LES RÉSUMÉS DU COURS
j'intègre

Gabriel Baudrand

Mathématiques :

résumés du cours
ECE 1re et 2e années

- Cours
- Exemples
- Applications
- Conseils



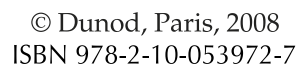
Mathématiques :

résumés du cours
ECE 1re et 2e année

Gabriel Baudrand

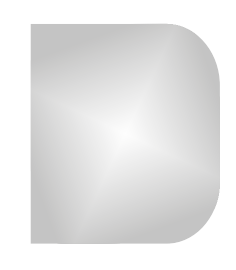
Professeur agrégé de mathématiques en classes préparatoires au lycée Madeleine Michelis (Amiens)



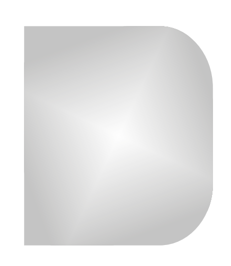


ISBN 978-2-10-053972-7

Table des matières



Mode d'emploi



Ce livre contient l'intégralité du cours de mathématiques pour les classes préparatoires ECE, première et deuxième années. Il intéressera aussi les étudiants en Licence de sciences économiques, et tous ceux qui désirent acquérir des connaissances élémentaires mais solides en analyse, algèbre linéaire, probabilités.

Quand on donne la définition d'un mot, celui-ci est imprimé en **gras**.

Les résultats essentiels sont encadrés.

Des éléments pour la démonstration d'un résultat sont donnés quand celle-ci utilise des techniques significatives et utiles pour la résolution des exercices. Ces éléments demandent au lecteur une participation active (rédiger complètement, faire les calculs omis), qui est la clé des progrès en mathématiques.

Les notions nouvelles sont illustrées par des exemples. Ceux-ci sont signalés en tant que que tels, ou par un liseré en marge gauche. Ils sont inspirés par des exercices très classiques ou provenant des annales de concours. Ils sont plus nombreux quand une grande variété de situations se présente.

Dans le même esprit, un certain nombre d'applications sont données. Elles ne font pas partie du cours, mais elles en sont le prolongement naturel, et inspirent de nombreux exercices d'annales. Ces caractéristiques sont signalées à chaque fois qu'il est nécessaire.

Sur fond grisé vous trouverez des conseils d'ordre pédagogique : écueils à éviter, erreurs à ne pas commettre, conseils de rédaction, remarques utiles à la mémorisation.

Quelques indications pour les différentes sections de ce livre

L'introduction expose les connaissances et techniques de base demandées par le programme. S'y ajoutent des considérations qui ne sont pas explicitement demandées, mais néanmoins indispensables : les éléments de logique aideront le lecteur à raisonner juste, ce qui aidera à une meilleure compréhension du cours. Les rappels sur les équations, inéquations, inégalités visent à consolider des acquis des classes antérieures et qui prennent maintenant toute leur importance.

En ce qui concerne l'analyse, la totalité du programme de terminale ES est reprise et bien sûr complétée. Les points les plus délicats du programme (recherche de limites, suites récurrentes, séries, intégrales généralisées) sont exposés progressivement et illustrés par de nombreux exemples.

Pour l'algèbre linéaire, la difficulté est d'une part technique (recherche des valeurs propres et vecteurs propres), et d'autre part théorique (utilisation des théorèmes abstraits du cours dans des situations diverses). On s'est efforcé de bien cerner les difficultés et ici aussi de donner suffisamment de variété dans les exemples.

En probabilités, on a choisi de traiter dans trois chapitres différents les problèmes concernant les variables aléatoires finies, discrètes, à densité. Cela oblige à quelques répétitions, mais les techniques différentes mises en œuvre (respectivement sommes finies, séries, calcul intégral) justifient une telle démarche. On a privilégié ici les démonstrations des résultats du cours, ou des applications les plus typiques, car leur maitrise est essentielle pour la résolution des exercices.

S'y ajoute un chapitre sur l'algorithmique : on y trouvera les éléments du langage PASCAL à connaître, et quelques programmes emblématiques.

Conformément au programme, les algorithmes (rédigés en PASCAL) viennent illustrer le cours. Ils sont encadrés par un liséré pointillé.

Pour ce qui concerne la répartition du travail sur les deux années de classe préparatoire, devraient être maitrisés en fin de première année :

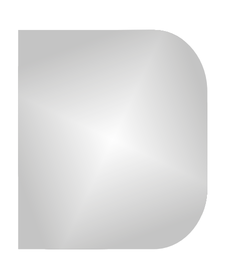
- l'introduction ;
- le chapitre 1, sauf § 1.1.5 ;
- le chapitre 2 sauf § 2.3 : notion de point fixe, et § 2.4 : critères de convergence et séries de Riemann ;
- le chapitre 3 : § 3.1 et 3.2, sauf sommes de Riemann et formules de Taylor ;
- le chapitre 4 ;

- les chapitres 7 et 8 (uniquement loi d'un couple, lois marginales et indépendance de deux v.a en ce qui concerne l'étude simultanée de plusieurs v.a).
- Pour ce qui concerne l'algorithmique, l'ensemble du programme est traité tout au long de la formation, à l'exception des algorithmes de gestion des listes, et tout ce qui concerne les v.a à densité et l'estimation, qui seront traités en deuxième année.

Dans le texte, les renvois commencent toujours par le numéro du chapitre (§ 2.3 renvoie au chapitre 2 paragraphe 3).



Introduction
Techniques de base



1. Ensembles, applications

1.1 Vocabulaire de la théorie des ensembles

$x \in E$: « x est élément de E », ou « x appartient à E ».

On ne cherche pas à définir les notions primitives d'**élément,** d'**appartenance**, d'**ensemble**.

On peut distinguer deux façons de définir un ensemble :

• Par extension : on donne la liste des éléments de l'ensemble. On notera en particulier, avec $n \in \mathbf{N}$:

$$[\![0, n]\!] = \{0 \ ; \ \dots \ ; n\}$$

• Par compréhension : on donne une propriété caractéristique P des éléments de l'ensemble. L'élément x appartient à l'ensemble E si, et seulement si, il vérifie la propriété P, ce que l'on note $P(x)$. Par exemple, a, b étant deux réels :

$$[a, b] = \{x \mid x \in \mathbf{R} \ ; a \leqslant x \leqslant b\}$$

Ici la propriété $P(x)$ est : « $x \in \mathbf{R}$ et $a \leqslant x \leqslant b$ ».

On rencontre des variantes de notation :

$$[a, b] = \{x \in \mathbf{R} \mid a \leqslant x \leqslant b\} = \{x \in \mathbf{R} \ ; a \leqslant x \leqslant b\} \dots$$

Certains ensembles ont des notations réservées :

$\varnothing$: l'ensemble vide (il ne contient aucun élément).

$\mathbf{N}$: l'ensemble des entiers naturels. $\mathbf{N} = \{0 \ ; 1 \ ; 2 \ ; \dots\}$.

$\mathbf{N}^*$: l'ensemble des entiers naturels non nuls.

$\mathbf{Z}$: l'ensemble des entiers relatifs.

$\mathbf{Q}$: l'ensemble des nombres rationnels.

R : l'ensemble des nombres réels.

$\mathbf{R}^+$: l'ensemble des nombres réels positifs ou nuls.

On définit de même $\mathbf{R}^*$, $\mathbf{R}^{+*}$, $\mathbf{R}^-$ ...

A, B, E ***étant des ensembles, on définit*** **:**

Relation d'inclusion. On note $A \subset E$ (lire « A est inclus dans E », ou « A est une partie de E », ou « A est un sous-ensemble de E ») si et seulement si tout élément de A est élément de E.

On note aussi $E \supset A$ (« E contient A »).

- Pour tout ensemble E, on a l'inclusion $\varnothing \subset E$.
- $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}$.

Réunion de deux ensembles. On note $A \cup B$ (lire « A union B ») l'ensemble ainsi défini :

$$A \cup B = \{x \mid x \in A \text{ ou } x \in B\}$$

Intersection de deux ensembles. On note $A \cap B$ (lire « A inter B ») l'ensemble ainsi défini :

$$A \cap B = \{x \mid x \in A \text{ et } x \in B\}$$

Généralisation : avec I un ensemble d'indices :

$$\bigcup_{i \in I} A_i = \{x \text{ ; il existe } i \in I \text{ tel que } x \in A_i\}$$

$$\bigcap_{i \in I} A_i = \{x \text{ ; pour tout } i \in I, x \in A_i\}$$

Complémentaire d'un ensemble dans un ensemble. Soit $A \subset E$. Le complémentaire de A dans E est l'ensemble des éléments de E qui n'appartiennent pas à A. On le note $E \setminus A$, ou, s'il n'y a pas d'ambiguïté sur l'ensemble E de référence, $\overline{A}$ (lire « A barre »).

Produit cartésien de deux ensembles. Le produit cartésien $A \times B$ est l'ensembles des couples $(a\ ;\ b)$ avec $a \in A$ et $b \in B$:

$$A \times B = \{(a\ ;\ b) \mid a \in A \text{ et } b \in B\}$$

On définit de même les produits cartésiens $A \times B \times C, \ldots$, et A^n :

$$A^n = \{(a_1\ ;\ \cdots\ ;\ a_n) \mid a_1 \in A\ ;\ \cdots\ ;\ a_n \in A\}$$

A^n est l'ensemble des suites à n éléments de A, ou ensemble des n-listes d'éléments de A ($n \in \mathbf{N}^*$).

Ensemble des parties de E. On note $\mathcal{P}(E)$ l'ensemble de toutes les parties de E :

$$\mathcal{P}(E) = \{A\ ;\ A \subset E\}$$

• Soit A, B, C des sous-ensembles de l'ensemble de référence E. On notera les règles de calculs :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \quad ; \quad \overline{A \cap B} = \overline{A} \cup \overline{B}$$

$$A \cup \varnothing = A \ ; \quad A \cap \varnothing = \varnothing \ ; \quad A \cup E = E \ ; \quad A \cap E = A$$

Règles de calcul qui se généralisent, par exemple :

$$B \cup \left(\bigcap_{i \in I} A_i \right) = \bigcap_{i \in I} (B \cup A_i)$$

• Remarquez que

$$A \subset B \Leftrightarrow A \cup B = B \Leftrightarrow A \cap B = A$$

1.2 Fonctions et applications

Définitions

• Une **fonction** f est définie par la donnée d'un ensemble de départ E, d'un ensemble d'arrivée F, et d'une relation qui à un élément de E associe au plus un élément de F. Notation :

$$f : E \to F, x \mapsto y = f(x)$$

• Si on a $y = f(x)$, on dit que y est l'**image** de x par f, et que x est un **antécédent** de y par f.

• Une **application** de E dans F est une fonction de E dans F telle que chaque élément de E admette une image. On note alors $f(E)$ l'ensemble $\{f(x) \ ; x \in E\}$.

• Soit $f : E \to F$ et $g : F \to G$ deux applications. La **composée** $g \circ f$ est l'application $g \circ f : E \to G, x \mapsto g \circ f(x) = g(f(x))$.

• On dit qu'une application $f : E \to F$ est :

– une **injection**, ou que f est **injective**, ssi tout élément de F admet au plus un antécédent : $f(x) = f\left(x'\right) \Rightarrow x = x'$.

– une **surjection**, ou que f est **surjective** ssi tout élément de F admet au moins un antécédent : pour tout $y \in F$, il existe $x \in E$ tel que $y = f(x)$.

– une **bijection**, ou que f est **bijective** ssi tout élément de F admet exactement un antécédent dans E : pour tout $y \in F$, il existe $x \in E$, x unique, tel que $y = f(x)$. On parle alors de bijection de E **sur** F.

Exemple important

Soit E un ensemble (non vide). L'application

$$\mathrm{Id}_E : E \to E, x \mapsto \mathrm{Id}_E(x) = x$$

est appelée l'**application identité** de E. (C'est d'ailleurs une bijection.)

Propriétés

Proposition 1. Une application est bijective ssi elle est injective et surjective.

Proposition 2. Soit f une bijection de E sur F. L'application, notée f^{-1} de F dans E qui à tout élément y de F associe l'unique élément x de E tel que $f(x) = y$ est une bijection de F sur E, appelée **bijection réciproque** de f:

$$f^{-1} : F \to E, y \mapsto f^{-1}(y) = x \text{ tel que } f(x) = y$$

Proposition 3. Une application f de E dans F est bijective ssi il existe une application g de F dans E, telle que $f \circ g = \mathrm{Id}_F$ et $g \circ f = \mathrm{Id}_E$.

On a alors $g = f^{-1}$.

• Il faut bien comprendre que l'ensemble de départ et d'arrivée sont essentiels dans la définition de l'application ou de la fonction f. Ainsi, les applications

$$f_1 : \mathbf{R} \to \mathbf{R}, x \mapsto x^2 \;;\quad f_2 : \mathbf{R} \to \mathbf{R}^+, x \mapsto x^2 \;;$$
$$f_3 : \mathbf{R}^+ \to \mathbf{R}^+, x \mapsto x^2$$

sont différentes. f_1 n'est ni injective ni surjective (les nombres négatifs n'ont pas d'antécédent par f_1, les nombres positifs en ont deux), f_2 est surjective mais pas injective, f_3 est bijective.

• La proposition 2 est essentielle, elle permet de définir de « nouvelles » fonctions.

– La bijection réciproque de f_3 est la fonction racine carrée.

– Le logarithme néperien est une bijection de $\mathbf{R}^{+*}$ sur $\mathbf{R}$. Sa bijection réciproque est la fonction exponentielle. La proposition 3 donne alors: Pour tout réel positif x, $e^{\ln x} = x$; pour tout réel y, $\ln(e^y) = y$.

2. Notions de logique

2.1 Généralités

Une **propriété** est une affirmation dont la valeur de vérité – vrai (V) ou faux (F) – peut dépendre de un ou plusieurs arguments, numériques ou autres. On notera $P(x)$ si la valeur de vérité de la proposition P dépend de la valeur de l'argument (ou variable) x. On dit alors que x est une **variable libre** pour la propriété P.

> x étant un nombre entier, la propriété $P(x)$: « x est un nombre premier » est vraie si $x = 2$, fausse si $x = 4$.

2.2 Quantificateurs

Définitions

Soit $P(x)$ une propriété, avec x appartenant à un ensemble de référence E.

- **Quantificateur existentiel.** La propriété

$$\exists x \in E, P(x)$$

est vraie si, et seulement si, il existe x appartenant à E tel que la propriété $P(x)$ soit vraie. On lit « il existe x appartenant à E tel que $P(x)$ », ou « pour quelque x appartenant à E, $P(x)$ ».

- **Quantificateur universel.** La propriété

$$\forall x \in E, P(x)$$

est vraie si, et seulement si, pour tout x appartenant à E, la propriété $P(x)$ est vraie. Lire « quel que soit x appartenant à E, $P(x)$. »

Exemples

L'ensemble de référence est $\mathbf{N}$. Soit les propriétés

$$P_1 : \forall x \in \mathbf{N}, x \geqslant 0 \quad ; \quad P_2(y) : \forall x \in \mathbf{N}, x \geqslant y \ ;$$
$$P_3 : \forall x \in \mathbf{N}, \exists y \in \mathbf{N}, x < y \quad ; \quad P_4 : \exists y \in \mathbf{N}, \forall x \in \mathbf{N}, x < y.$$

P_1 est vraie (tout entier naturel est supérieur ou égal à 0).

$P_2(y)$ est vraie si $y = 0$, fausse dans tous les autres cas.

P_3 est vraie : tout entier naturel admet un entier qui lui est supérieur.

P_4 est fausse : il n'existe pas d'entier naturel supérieur à tous les autres. À noter que l'ordre des quantificateurs a de l'importance.

Dans P_3, ni x ni y ne sont des variables libres. P_3 est une propriété de $\mathbf{N}$, pas de x, ni de y, qui sont ici des **variables muettes**. On pourrait écrire P_3 sous la forme : $\forall a \in \mathbf{N}, \exists b \in \mathbf{N}, a < b$.

2.3 Opérateurs logiques

Définitions

Soit P, Q, deux propriétés.

• La propriété P **ou** Q est vraie si une des deux propriétés (ou les deux) est vraie, fausse si P et Q sont fausses.

• La propriété P **et** Q est vraie si les deux propriétés sont vraies (simultanément), fausse si une deux propriétés (ou les deux) est fausse.

• La propriété **non** P est vraie si P est fausse, fausse si P est vraie.

• La propriété « **si** P, **alors** Q » est fausse si P est vraie et Q fausse, vraie dans tous les autres cas . On dit aussi :

« P est une condition suffisante de Q »,

« Pour Q, il suffit que P »,

« Q est une condition nécessaire de P »,

« Pour P, il faut que Q »,

« P seulement si Q »,

« P implique Q », et on note $P \Rightarrow Q$.

• La propriété « P **si et seulement si** Q » est vraie si P et Q ont même valeur de vérité, fausse sinon. On dit aussi :

« P est une condition nécessaire et suffisante de Q »,

« P et Q sont équivalentes »,

« pour que Q, il faut et il suffit que P »,

et on note $P \Leftrightarrow Q$.

On écrit couramment en abrégé « ssi » pour « si et seulement si ».

Ces définitions sont synthétisées dans les **tables de vérité** :

P	Q	**P ou Q**	**P et Q**	$P \Rightarrow Q$	$P \Leftrightarrow Q$
V	V	V	V	V	V
V	F	V	F	F	F
F	V	V	F	V	F
F	F	F	F	V	V

Règles de calcul

Les propriétés suivantes sont équivalentes :

non (P ou Q) et (non P) et (non Q) ;
non (P et Q) et (non P) ou (non Q) ;
non ($\exists x, P(x)$) et $\forall x$, non ($P(x)$) ;
non ($\forall x, P(x)$) et $\exists x$, non ($P(x)$) ;
non ($P \Rightarrow Q$) et Q et non (P).

Les règles de calcul ci-dessus sont utiles pour montrer qu'une propriété est fausse. Par exemple, pour montrer qu'une propriété universelle $(\forall x, P(x))$ est fausse, il suffit de donner un **contre-exemple**, c'est-à-dire une valeur de x telle que $P(x)$ est fausse.

Utilisation

La plupart des théorèmes et propositions du cours se présentent comme des implications (vraies !) « P implique Q », ou comme des équivalences. Le vocabulaire impliqué est d'usage constant et doit être bien compris. En particulier, on notera qu'une condition nécessaire peut ne pas être suffisante, et qu'une condition suffisante peut ne pas être nécessaire :

> (§ 2.4) Pour qu'une série soit convergente, il est nécessaire, mais pas suffisant, que son terme général tende vers 0. En d'autres termes, si le terme général ne tend pas vers 0, alors la série ne converge pas, mais si le terme général tend vers 0, la série peut ne pas converger.
>
> (§ 6.2) Pour qu'une matrice soit diagonalisable, il est suffisant, mais pas nécessaire, qu'elle soit symétrique.

À noter le lien avec le vocabulaire des ensembles. Avec A, B inclus dans l'ensemble de référence E, si on a

$A = \{x|P(x)\}$; $B = \{x \mid Q(x)\}$, alors

$A \cup B = \{x \mid P(x) \text{ ou } Q(x)\}$; $A \cap B = \{x|P(x) \text{ et } Q(x)\}$

$\overline{A} = \{x \mid \text{non}(P(x))\}$

$A \subset B$ si et seulement si $P(x) \Rightarrow Q(x)$.

• Ce qu'on a appelé ici « propriétés » correspond à ce qu'on appelle en langage PASCAL les **variables booléennes**, dont le contenu est TRUE (vrai) ou FALSE (faux). Les opérateurs logiques OR, AND, NOT correspondent aux opérateurs sur les propriétés vus ici. Mais attention l'instruction IF ... THEN... n'est pas un opérateur logique : THEN est suivie d'une instruction, pas d'une variable booléenne.

2.4 Quelques méthodes de raisonnement

Raisonnement par récurrence

Soit à établir qu'une propriété $P(n)$ est vraie pour tout $n \in \mathbf{N}$.

• On établit que $P(0)$ est vraie (**initialisation**).

• On suppose qu'il existe $n \in \mathbf{N}$ tel que $P(n)$ est vraie (**hypothèse de récurrence**). On montre alors que $P(n+1)$ est vraie (**hérédité**).

• On conclut alors, d'après le **principe de récurrence** :

$$\forall n \in \mathbf{N}, P(n).$$

Le raisonnement par récurrence doit être considéré comme un véritable guide de rédaction. Celui-ci doit être suivi scrupuleusement et rédigé soigneusement. Cela n'empêche pas que le cas échéant la rédaction puisse être rapide et synthétique.

Voici quelques situations typiques où on fait un raisonnement par récurrence qui ne présente aucune difficulté :

• (§ 4.2.4) Soit $A \in \mathbf{M}_3(\mathbf{R})$, $X_n \in \mathbf{M}_{3,1}(\mathbf{R})$ telles que $\forall n \in \mathbf{N}$, $X_{n+1} = AX_n$. On montre alors par récurrence :

$$\forall n \in \mathbf{N}, \quad X_n = A^n X_n$$

• (§ 2.3) Soit $(u_n)_{n\in\mathbf{N}}$ une suite telle que $\forall n \in \mathbf{N}, u_{n+1} = f(u_n)$, et $u_0 = a$, avec a tel que $f(a) = a$. On montre alors par récurrence :

$$\forall n \in \mathbf{N} \ \ u_n = a$$

• Pour établir l'hérédité, il faut souvent utiliser une idée ou une propriété mise en évidence dans une question précédente. C'est le cas pour le premier exemple ci-dessus (la propriété qui permet d'établir l'hérédité est $X_{n+1} = AX_n$), et d'une manière tout à fait typique pour l'étude de suites récurrentes grâce à la formule des accroissements finis (cf § 2.3.3).

• Le raisonnement par récurrence est susceptible de nombreuses variations : l'initialisation peut être faite avec $n = 1$. L'hérédité permet de passer de $n - 1$ à n ($n \in \mathbf{N}^*$)...

Parfois l'initialisation devra porter sur les propriétés $P(0)$, $P(1)$, $P(2)$, par exemple, et pour obtenir l'hérédité on supposera qu'il existe $n \in \mathbf{N}$ tel que $P(n)$, $P(n+1)$, $P(n+2)$ sont vraies (récurrence sur plusieurs générations). Ou bien on supposera qu'il existe $n \in \mathbf{N}$ tel que, pour tout $k \in \{0 \ ; \ \cdots \ ; n\}$, $P(k)$ est vraie (récurrence forte).

Raisonnement par contraposée

La contraposée de la propriété $P \Rightarrow Q$ est la propriété non $Q \Rightarrow$ non P. Elles sont logiquement équivalentes, et pour établir une implication, il peut être plus commode d'établir sa contraposée.

Pour montrer qu'un polynôme de degré $n \geqslant 1$ admet au plus n racines, on démontre la contraposée : un polynôme admettant plus de n racines n'est pas de degré n. Voir le § 6 de cette introduction.

Ne pas confondre contraposée et réciproque : la réciproque de la propriété $P \Rightarrow Q$ est la propriété $Q \Rightarrow P$: la réciproque d'une implication vraie peut être fausse.

L'implication « la série Σu_n converge $\Rightarrow \lim_{n\to+\infty} u_n = 0$ » est vraie.

Sa contraposée est vraie, mais sa réciproque est fausse : la suite $\left(\frac{1}{n}\right)_{n\in\mathbf{N}^*}$ converge vers 0, et la série $\sum_{n\in\mathbf{N}} \frac{1}{n}$ diverge (voir § 1.4.2).

Raisonnement par l'absurde

Pour montrer qu'une propriété P est vraie, on suppose qu'elle est fausse, et on aboutit à une contradiction. On conclut alors que P est vraie.

Deux matrices (voir chap. 4) A et B sont données, B est non nulle et $AB = 0$. On montre par l'absurde que A n'est pas inversible : pour cela on suppose que A est inversible, et de $AB = 0$ on tire alors $A^{-1}AB = A^{-1}0$, donc $IB = 0, B = 0$. Or $B \neq 0$: contradiction, A n'est pas inversible.

Mentionnons aussi le **raisonnement par équivalence** (utilisé dans la résolution des équations) : on montre qu'une propriété est équivalente à une propriété vraie, plus simple. Le **raisonnement par analyse et synthèse** consiste à trouver des conditions nécessaires à l'existence d'un objet (la solution d'une équation par exemple), puis à vérifier si ces conditions nécessaires sont suffisantes.

3. Signes Σ, Π

3.1 Définitions

Soit I un sous-ensemble **fini** de $\mathbf{N}$ ou de $\mathbf{N} \times \mathbf{N}$. Les symboles

$$\sum_{i\in I} x_i \ ; \qquad \prod_{i\in I} x_i$$

désignent respectivement la somme et le produit de tous les nombres réels x_i, avec i appartenant à I.

Cas particuliers, d'utilisation très fréquente :

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n \quad ; \quad \sum_{i=0}^{n} x_i = x_0 + x_1 + \cdots + x_n$$

Lire « sigma de i égal 1 à n des x indice i »... Notations analogues pour le produit (lire « produit de i égal 1 à n... »).

3.2 Règles de calcul

- Avec I fini inclus dans $\mathbf{N}$, et a constante indépendante de i, on a

$$\sum_{i\in I}(x_i+y_i)=\sum_{i\in I}x_i+\sum_{i\in I}y_i$$

$$\sum_{i\in I}ax_i=a\sum_{i\in I}x_i$$

- Avec $n\in\mathbf{N}^*,\ p\in\mathbf{N}, p\leqslant n$ (et a constante indépendante de i) :

$$\sum_{i=1}^{n}a=na\ ;\quad \sum_{i=0}^{n}a=(n+1)\,a\ ;\quad \sum_{i=p}^{n}a=(n-p+1)\,a$$

Démonstration. La première propriété est une conséquence de la commutativité de l'addition. La deuxième propriété est une mise en facteur commun. Pour les propriété suivantes, on compte combien la somme contient de termes tous égaux à a.

- Si les x_i sont des réels positifs, les y_i des réels quelconques :

$$\ln\left(\prod_{i\in I}x_i\right)=\sum_{i\in I}\ln x_i\quad ;\quad \exp\left(\sum_{i\in I}y_i\right)=\prod_{i\in I}\exp(y_i)$$

- Si I est une partie finie de $\mathbf{N}\times\mathbf{N}$, il s'agit en fait d'une « somme double », qu'on décompose en somme de sommes :

$$\sum_{(i,j)\in I}x_{i,j}=\sum_{\{i\ ;\ \exists j,(i,j)\in I\}}\left(\sum_{\{j\ ;\ (i,j)\in I\}}x_{i,j}\right)=\sum_{\{j\ ;\ \exists i,(i,j)\in I\}}\left(\sum_{\{i\ ;\ (i,j)\in I\}}x_{i,j}\right)$$

En particulier, on a, avec $I=\{(i,j)\ ;\ 1\leqslant i\leqslant n \text{ et } 1\leqslant j\leqslant i\}$:

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{i}x_{i,j}\right)=\sum_{j=1}^{n}\left(\sum_{i=j}^{n}x_{i,j}\right)$$

- Les règles de calcul ci-dessus sont les seules à retenir, et à utiliser. Vous prendrez garde à ne pas en inventer d'autres, qui ont de bonnes chances d'être fausses. Exemple : bien se persuader qu'en général

$$\sum_{i=1}^{n}x_iy_i\neq\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)$$

Avec $n=2$, en effet, on aura $x_1y_1+x_2y_2\neq(x_1+x_2)(y_1+y_2)$.

• Par contre, il est vrai que

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n} x_i y_j\right) = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} y_j\right) = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=1}^{n} y_j\right)$$

(Mise en facteur de x_i dans la somme $\sum_{j=1}^{n}$, puis mise en facteur de $\sum_{j=1}^{n}$.)

• Si vous êtes « bloqué(e) » dans un calcul comportant un $\sum$, une possibilité est d'expliciter le $\sum$ en question, sur le modèle

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

On écrit les deux ou trois premiers termes de la somme, puis le dernier. Mais il faut faire attention alors que si $n = 1$, la somme ne comporte quand même qu'un seul terme, le terme x_1.

• Prenez bien garde au statut des variables en présence :

Dans la somme $\sum_{i=1}^{n} x_i$, i est une variable muette : la valeur de i n'intervient pas dans la valeur de la somme. On a par exemple $\sum_{i=1}^{n} x_i = \sum_{k=1}^{n} x_k$.

Dans cette même somme, n est une variable libre : la valeur de la somme dépend *a priori* de la valeur de n. N'écrivez donc pas des formules du type $\sum_{i=1}^{n} x_i = f(i)$ qui n'ont aucune chance d'être vraies...

3.3 Sommes remarquables

On retient la valeur des sommes suivantes. n est un entier naturel non nul, x un nombre réel.

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}$$

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} \quad \text{si } x \neq 1$$

Voir le § 2.2.1 (suites arithmétiques) pour la démonstration du premier résultat, et le paragraphe § 2.2.2 pour celle du dernier.

Les deuxième et troisième résultats (somme des carrés, somme des cubes) se démontrent classiquement par récurrence.

4. Dénombrement — Formule du binôme

Un ensemble non vide E est dit **fini** ssi il existe $n \in \mathbf{N}$ tel qu'il existe une bijection de E sur $[\![1\ ;\ n]\!]$. n est alors le **cardinal** de E.

On note $\operatorname{Card} E = n$. L'ensemble vide est fini, de cardinal 0.

Le cardinal d'un ensemble fini est simplement le nombre de ses éléments.

Un ensemble est dit **dénombrable** ssi il existe une bijection de **N** sur cet ensemble. Attention, les problèmes de dénombrement concerne les ensembles finis !

4.1 Factorielle d'un nombre entier

Définition

Pour n appartenant à **N**, on définit par récurrence :

$$0! = 1\ ;\ \text{si } n \geqslant 1, n! = n \times (n-1)!$$

Lire « factorielle n ». D'après la définition, on a

$$1! = 1\ ;\ 2! = 2 \times 1 = 2\ ;\ 3! = 3 \times 2 \times 1 = 6\ ;\ 4! = 24\ ;\ \ldots$$

De façon générale, pour $n \geqslant 1$: $n! = n \times (n-1) \times \cdots \times 1$

La programmation de $n!$ en langage PASCAL peut se faire de façon itérative :

```
function faciter(n :integer) :real ;
var k:integer; fac:real;
begin
  fac:=1;
  for k:=1 to n do fac:=fac*k;
  faciter:=fac;
end;
```

La programmation récursive est ici préférable :

```
function facrec(n :real):real;
begin
  if n=0 then facrec:=1 else
  facrec:=n*facrec(n-1);
end;
```

On rencontre souvent des simplifications du type :

$$\frac{(n+1)!}{n!} = n+1 \quad ; \quad \frac{n!}{(n-k)!} = n\,(n-1)\ldots(n-k+1) \ ; \ \ldots$$

4.2 Formules élémentaires

Nombre de termes

p et n étant des entiers naturels tels que $p \leqslant n$, de p à n il y a $n-p+1$ nombres entiers.

Ainsi, de 1 à 100, il y a 100 nombres entiers, et non 99. De 100 à 200 il y a $200 - 100 + 1 = 101$ nombres entiers.

Nombre de suites finies

- Soit $n, k \in N$. Le nombre de suites à k éléments d'un ensemble à n éléments est égal à n^k.
- Soit $n, k \in \mathbf{N}$ tels que $1 \leqslant k \leqslant n$. Le nombre de suites à k éléments distincts d'un ensemble à n éléments est égal à

$$n\,(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

- Soit $n \in \mathbf{N}$; le nombre de suites à n éléments distincts de l'ensemble E à n éléments, ou **permutations** de E, est égal à $n!$.

On obtient ces formules à l'aide de représentations arborescentes.

La première est un cas particulier de la formule

$\text{Card}(A_1 \times A_2 \times \cdots \times A_k) = \text{Card}(A_1) \cdot \text{Card}(A_2) \cdot \ldots \cdot \text{Card}(A_k)$,

avec tous les A_i égaux, de cardinal n.

Elle est valable même si n ou k sont nuls, si on considère la « suite vide » (suite à 0 élément), et avec la convention $0^0 = 1$.

De même, en adoptant l'écriture avec des factorielles, la deuxième formule reste valable même si k ou n sont nuls. Notez que le produit

$$n\,(n-1)\ldots(n-k+1)$$

comporte k facteurs (autant que de nombres de 0 à $k-1$).

La troisième formule est un cas particulier important de la deuxième, avec $k = n$. Une permutation d'un ensemble à n éléments peut être vue comme une manière d'écrire dans un certain ordre les éléments de cet ensemble, ou comme une bijection de $[\![1, n]\!]$ sur cet ensemble.

Cardinal de la réunion de deux ensembles

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)\ ;$$
$$\text{Si } A \cap B = \varnothing, \text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B).$$

Formule qui se généralise à 3, 4, n ensembles sur le modèle de la formule du crible, voir § 7.1.2, où on remplacera les « P » par des « Card ».

4.3 Nombre de parties d'un ensemble

Théorème

• Soit $k, n \in \mathbf{N}$, $0 \leqslant k \leqslant n$. Le nombre de parties à k éléments d'un ensemble à n éléments est égal à :

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

• Le nombre de parties d'un ensemble à n éléments est égal à 2^n:

$$\text{Si Card}(E) = n, \text{ alors Card}(\mathcal{P}(E)) = 2^n$$

En effet le nombre de suites à k éléments distincts d'un ensemble à n éléments est égal à $\frac{n!}{(n-k)!}$. Mais chaque partie à k éléments de cet ensemble à n éléments est représentée par $k!$ permutations distinctes, d'où le premier résultat. On en déduit le deuxième résultat en utilisant la formule du binôme, voir § 2.4.4.

• $\binom{n}{k}$ se lit « k parmi n ». C'est le nombre de manières de choisir k éléments parmi n, quand on ne tient pas compte de l'ordre du choix.

• Les nombres $\binom{n}{k}$ sont appelés coefficients binomiaux, voir § 2.4.4.

Après simplifications, quand $1 \leqslant k \leqslant n$, on peut écrire

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$$

Le numérateur et le dénominateur comportent chacun k facteurs.

Vous utiliserez cette technique, et la formule $\binom{n}{n-k} = \binom{n}{k}$, pour calculer des valeurs particulières, par exemple

$$\binom{10}{7} = \binom{10}{3} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 720$$

Propriétés de nombres $\binom{n}{k}$

Soit $n, k \in \mathbf{N}$, $0 \leqslant k \leqslant n-1$. Alors :

$$\binom{n}{0} = \binom{n}{n} = 1 \;;\quad \binom{n}{1} = \binom{n}{n-1} = n \;;\quad \binom{n}{n-k} = \binom{n}{k} \;;$$

(formule de Pascal) $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

Ces formules se démontrent en utilisant les factorielles, ou bien par des considérations de dénombrement : il est évident par exemple que $\binom{n}{0} = 1$. Pour la formule de Pascal, considérer les parties à $k+1$ éléments de l'ensemble $\{a_1, a_2, \ldots, a_n, b\}$: il y en a $\binom{n+1}{k+1}$, et celles qui contiennent b sont au nombre de $\binom{n}{k}$, celles qui ne contiennent pas b sont au nombre de $\binom{n}{k+1}$.

Elles permettent de construire de proche en proche le **triangle de Pascal**, ou figure en ligne n et colonne k le nombre $\binom{n}{k}$:

$n \backslash k$	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
6	1	6	15	20	15	6	1

Programmation du triangle de Pascal jusqu'à la ligne n : on utilise des boucles définies emboîtées. La variable d'entrée est `n`, celle de sortie est `c`, de type tableau.

```
 rogram  ascal ;
var n,i,j:integer;c:array[0..100,0..100]of integer;
BEGIN   readln(n) ;
{initialisation de la première colonne et de la diagonale}
for i :=0 to n do  begin c(0,0):=1; c(i,i):=1;
end;
{initialisation du reste du tableau par la formule de Pascal}
for i:=1 to n-1 do
   for j:=0 to i-1 do c(i+1,j+1):=c(i,j)+c(i,j+1);
```

```
{écriture}  for i:=0 to n do
                begin  for j:=0 to i do
                    write(c(i,j));writeln;   end;
END.
```

Comme autres propriétés des nombres $\binom{n}{k}$, mentionnons :

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} \ ; \quad \binom{n}{s}\binom{s}{k} = \binom{n}{k}\binom{n-k}{s-k} \ ;$$

$$\sum_{i=k}^{n}\binom{i}{k} = \binom{n+1}{k+1} \ ;$$

$$\sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k} = \binom{n+m}{p} \ ; \quad \sum_{k=0}^{n}\binom{n}{k}^2 = \binom{2n}{n}$$

Les deux premières égalités s'établissent très facilement avec les factorielles. La première est fréquemment utilisée.

La troisième se démontre par récurrence sur n: la propriété est vraie pour $n = 0$. On la suppose vraie pour n fixé dans $\mathbf{N}$; pour $k \in [\![0 \,;\, n]\!]$, on a alors

$$\sum_{i=k}^{n+1}\binom{i}{k} = \sum_{i=k}^{n}\binom{i}{k} + \binom{n+1}{k} = \binom{n+1}{k+1} + \binom{n+1}{k} = \binom{n+2}{k+1}$$

(hypothèse de récurrence, puis formule de Pascal). Pour $k = n + 1$, la propriété est aussi vraie, l'hérédité est ainsi complètement établie.

La quatrième égalité est connue sous le nom de **formule de Vandermonde**. Elle est facile à mémoriser si on pense à un exemple : avec $n = 8, m = 24, p = 5$, le nombre de mains de 5 cartes d'un jeu de 32 cartes (membre de droite de l'égalité) est égal à la somme des nombres de mains de 5 cartes comportant 0, 1, 2, 3, 4, 5 cœurs (membre de gauche ; en effet, le nombre de mains de 5 cartes comportant k cœurs est égal à $\binom{8}{k}\binom{24}{5-k}$: on choisit k cœurs parmi les 8 cœurs du jeu de 32 cartes, puis on complète avec $5 - k$ non-cœurs parmi 24 non-cœurs).

La cinquième formule est un cas particulier de la quatrième, avec :

$m = p = n$, en utilisant $\binom{n}{n-k} = \binom{n}{k}$.

4.4 Formule du binôme

On démontre classiquement par récurrence le

Théorème. Pour tout $a, b \in \mathbf{R}$, $n \in \mathbf{N}$ **:**

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Pour les premières valeurs de n, on obtient :

$$(a+b)^0 = 1$$

$$(a+b)^1 = a+b$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

Vous devez mémoriser, et savoir utiliser cette formule dans les deux sens (développement ou factorisation). Voici quelques remarques pour aider à cette mémorisation :

Dans le développement, les coefficients sont ceux de la ligne n du tableau de pascal. Le premier développement est fait suivant les puissances décroissantes de a, croissantes de b, et c'est le contraire pour le deuxième développement. Dans chaque terme, la somme des exposants est égale à n.

Quelques exemples d'utilisation

- Pour obtenir le développement de $(a-b)^n$, on écrit :

$$(a-b)^n = (a+(-b))^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}(-1)^k b^k$$

- Avec $a = b = 1$, on obtient

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

ce qui démontre que **le nombre de parties d'un ensemble à** n **éléments est égal à** 2^n.

• Avec $a = 1$, $b = -1$, on obtient :

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$$

ce qui montre que dans un ensemble à n éléments, le nombre de parties de cardinal pair et le nombre de parties de cardinal impair sont tous deux égaux, et donc égaux à 2^{n-1}.

• Voici un exercice de dénombrement où la formule du binôme est utilisée : si E est un ensemble de cardinal n, alors le nombre de couples (A, B) de parties de E telles que $A \cap B = \varnothing$ est égal à 3^n.

En effet, pour tout k appartenant à $[\![0, n]\!]$, le nombre de parties A à k éléments est $\binom{n}{k}$. Le choix de A étant fait, le nombre de parties B telles que $A \cap B = \varnothing$ est 2^{n-k}, car $A \cap B = \varnothing$ ssi $B \subset \overline{A}$, ssi B est une partie de $\overline{A}$, et le nombre de parties de $\overline{A}$ est égal à 2^{n-k}, $\overline{A}$ étant de cardinal $n - k$. En faisant varier k de 0 à n, on obtient un nombre de couples (A, B) qui conviennent égal à

$$\sum_{k=0}^{n} \binom{n}{k} 2^{n-k} = \sum_{k=0}^{n} \binom{n}{k} 2^{n-k} 1^k = (2+1)^n = 3^n$$

5. Équations, inéquations

Résoudre une équation d'inconnue $x \in \mathbf{R}$, c'est déterminer l'ensemble des nombres réels par lesquels on peut remplacer l'inconnue x de façon à obtenir une égalité vraie.

Définitions analogues pour un inéquation, un système d'équations.

5.1 Problèmes du premier degré

Théorème (équation du premier degré).
Soit $a, b \in \mathbf{R}$, et soit $S = \{x \in \mathbf{R}\,;\, ax + b = 0\}$. Alors

si $a \neq 0$, $S = \left\{-\dfrac{b}{a}\right\}$;

si $a = 0$ et $b \neq 0$, $S = \varnothing$; si $a = 0$ et $b = 0$, $S = \mathbf{R}$.

En particulier, l'équation $ax + b = 0$ admet une solution unique ssi $a \neq 0$.

Cette remarque est tout à fait essentielle et trouve sa généralisation dans les systèmes d'équations linéaires, voir § 4.1.

Vous serez particulièrement vigilant devant une équation telle que $ax = 0$: si $a = 0$, l'équation devient $0 \times x = 0$, qui a une infinité de solutions. Sinon l'équation devient $x = 0$, qui a une solution unique.

Signe de $ax + b$, Pour $a \neq 0$:

x	$-\infty$	$-b/a$	$+\infty$
$ax + b$	signe de $(-a)$	0	signe de a

5.2 Problèmes du second degré

Soit a, b, c des nombres réels, avec $a \neq 0$, et soit le polynôme du second degré $P(x) = ax^2 + bx + c$. Tous les résultats à connaître découlent de ce calcul :

$$P(x) = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) = a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right]$$
$$= a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right]$$

Soit $\Delta = b^2 - 4ac$ (discriminant).

• Si $\Delta < 0$, $P(x)$ est de la forme $a\left(B^2 + C\right)$ avec $C > 0$: $P(x)$ ne s'annule pas, ne se factorise pas dans $\mathbf{R}$, est toujours du signe de a.

• Si $\Delta = 0$, alors $P(x) = a\left(x + \frac{b}{2a}\right)^2$, admet $-\frac{b}{2a}$ pour racine double, se factorise dans $\mathbf{R}$, est du signe de a en dehors de la racine.

• Si $\Delta > 0$, $P(x)$ est de la forme $a\left(B^2 - C^2\right) = a(B - C)(B + C)$. On obtient

$$P(x) = a(x - x_1)(x - x_2)$$

avec $x_1 = \dfrac{-b - \sqrt{\Delta}}{2a}$; $x_2 = \dfrac{-b + \sqrt{\Delta}}{2a}$

$P(x)$ admet les deux racines distinctes x_1 et x_2, est factorisable dans $\mathbf{R}$, est « du signe de a à l'extérieur des racines », du signe de $-a$ à l'intérieur.

On voit donc que pour un polynôme du second degré, $P(x)$ est factorisable par $x - a$ ssi a est racine de $P(x)$. Ce résultat se généralise aux polynômes de degré supérieur, voir le § 6 de cette introduction.

• Ces formules de résolution sont relativement sophistiquées, et vous devriez éviter de les utiliser quand c'est possible, c'est-à-dire assez souvent :

– Équation $ax^2 + c = 0$: isoler x^2, ou factoriser quand c'est possible.
– Équation $ax^2 + bx = 0$: factoriser par x.
– « Racine évidente » : si $a + b + c = 0$, alors $ax^2 + bx + c$ admet 1 pour racine, on trouve l'autre en factorisant par $x - 1$; elle vaut $\frac{c}{a}$.

De même, si $a - b + c = 0$, les deux racines sont -1 (évidente) et $-\frac{c}{a}$.

• De façon générale, quand elles existent, la somme et le produit des racines de $ax^2 + bx + c$ valent respectivement $\frac{c}{a}$ et $-\frac{b}{a}$ (développer $a(x - x_1)(x - x_2) = ax^2 + bx + c$, puis identifier).

• Pour déterminer la position d'un nombre m par rapport aux racines de $P(x)$, il n'est pas nécessaire de déterminer celles-ci, il suffit de calculer $P(m)$, qui sera du signe de a si m est à l'extérieur des racines.

• Si l'équation est à coefficients entiers et b est pair, vous simplifierez par 2 l'expression des racines (si elles existent).

5.3 Autres équations et inéquations

Il n'y a pas de théorie générale et qui marcherait dans toutes les situations. Contentons-nous de donner quelques principes :

• Problèmes du type $P(x) = 0, P(x) \geqslant 0$, où P est un polynôme. On se ramène par factorisation à des problèmes de degré $\leqslant 2$, en utilisant le théorème de factorisation des polynômes, voir le § 6 de cette introduction.

• Problèmes du type $R_1(x) = R_2(x), R_1(x) \leqslant R_2(x)$, où R_1, R_2 sont des fractions rationnelles, ou fonctions rationnelles, c'est-à-dire des quotients de polynômes. On élimine les dénominateurs, en suivant les principes : $\frac{A}{B} = 0 \Leftrightarrow A = 0$ et $B \neq 0$; le signe de $\frac{A}{B}$ est déterminé par le signe de A et de B.

$$\frac{x^2}{x-1} \geqslant \frac{1}{x-1} \Leftrightarrow \frac{x^2-1}{x-1} \geqslant 0 \text{ (surtout pas équivalent à } x^2 \geqslant 1\text{ !)}$$
$$\Leftrightarrow \frac{(x-1)(x+1)}{x-1} \geqslant 0 \Leftrightarrow x+1 \geqslant 0 \text{ et } x \neq 1 \Leftrightarrow x \in]-1, +\infty[\setminus \{1\}$$

• Problèmes irrationnels, avec présence d'un (ou plusieurs) radicaux. On isole le radical, puis on élève au carré, mais attention aux équivalences :

$$\sqrt{A} = B \Leftrightarrow A = B^2 \text{ et } B \geqslant 0 \ ; \quad \sqrt{A} \leqslant B \Leftrightarrow 0 \leqslant A \leqslant B^2 \text{ et } B \geqslant 0$$

Résoudre $f(x) = x, f(x) \leqslant x$, avec $f(x) = \frac{x+1}{\sqrt{x^2+1}} - 1$.

$$f(x) = x \Leftrightarrow (x+1)\left(\frac{1}{\sqrt{x^2+1}} - 1\right) = 0$$

$$\Leftrightarrow x = -1 \text{ ou } \sqrt{x^2+1} = 1$$

$$\Leftrightarrow x = -1 \text{ ou } x^2 + 1 = 1 \Leftrightarrow x = -1 \text{ ou } x = 0$$

$$f(x) \leqslant x \Leftrightarrow (x+1)\left(\frac{1-\sqrt{x^2+1}}{\sqrt{x^2+1}}\right) \leqslant 0$$

$$\Leftrightarrow (x+1)\left(1-\sqrt{x^2+1}\right) \leqslant 0$$

$$1 - \sqrt{x^2+1} \leqslant 0 \Leftrightarrow \sqrt{x^2+1} \geqslant 1 \Leftrightarrow x^2 + 1 \geqslant 1 \Leftrightarrow x \in \mathbf{R}$$

Un tableau de signes conduit finalement à l'ensemble des solutions $[-1\,;\ +\infty[$.

• Autres types d'équations, comportant des fonctions logarithmes ou exponentielles.

— $e^{2x} - 3e^x + 2 = 0 \Leftrightarrow y = e^x$ et $y^2 - 3y + 2 = 0 \Leftrightarrow$

$$e^x = 1 \text{ ou } e^x = 2 \Leftrightarrow x = 0 \text{ ou } x = \ln 2$$

— $e^x - e^{-x} > 0 \Leftrightarrow e^x > e^{-x} \Leftrightarrow x > -x \Leftrightarrow x > 0$.

D'autres techniques sont parfois nécessaires, en particulier l'étude d'une fonction :

— Pour tout $x \in \mathbf{R}$, $e^x \geqslant x + 1$. En effet, avec $g(x) = e^x - x - 1$, on a $g'(x) = e^x - 1$, donc

x	$-\infty$		0		$+\infty$
$g'(x)$		$-$	0	$+$	
$g(x)$		$\searrow$	0	$\nearrow$	

d'où la conclusion (l'étude des limites est inutile).

— On montre de même : $\forall x > 0, \ln x \leqslant x - 1$.

Vous prendrez bien garde à l'énoncé ; la question :

« Résoudre l'équation $f(x) = 0$ »

est tout à fait différente de la question

« Montrer que l'équation $f(x) = 0$ admet une solution unique ».

Pour la première question la valeur explicite de la solution, ou des solutions, est demandée. Ce n'est pas le cas de la dernière question, et il ne faut donc pas chercher à exprimer cette solution.

Ainsi on peut prouver par des techniques d'analyse que l'équation

$$e^{-x} = x$$

admet une solution unique, et en donner une valeur approchée, mais il ne faut surtout pas chercher à en donner la valeur exacte !

6. Polynômes

6.1 Définitions

Un **polynôme**, ou **fonction polynôme**, est une application P de $\mathbf{R}$ dans $\mathbf{R}$ définie par :

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n$$

où les a_i sont des nombres réels.

Le plus grand entier i tel que $a_i \neq 0$ est appelé le **degré** de P. On note $i = d^\circ(P)$. Si tous les a_i sont nuls, P est le polynôme nul, on convient que son degré est $-\infty$.

On dit que le réel α est une **racine** de P ssi $P(\alpha) = 0$.

On dit que le polynôme P est **factorisable** (ou **divisible**) par le polynôme Q ssi il existe un polynôme R tel que $P(x) = Q(x)\,R(x)$.

6.2 Propriétés

Propriétés algébriques

• Si P et Q sont des polynômes, alors $P + Q$ et PQ sont des polynômes, et on a :

$$d^\circ(PQ) = d^\circ(P) + d^\circ(Q)\ ;\ d^\circ(P + Q) \leqslant \mathrm{Max}\,(d^\circ(\mathrm{P}), d^\circ(\mathrm{Q}))\ ;$$

Les résultats sur le degré sont valables même si un des polynômes est nul, avec la convention $-\infty + b = -\infty$.

• La composée de deux polynômes est un polynôme, et on a, avec P et Q non nuls : $d^\circ(P \circ Q) = d^\circ(P) \times d^\circ(Q)$.

• La dérivée d'un polynôme de degré n, $n \geqslant 1$, est un polynôme de degré $n - 1$.

Théorème de factorisation des polynômes

Le polynôme $P(x)$ est factorisable par $x - a$ ssi a est racine de P :

$$P(x) = (x - a)\,Q(x) \text{ avec Q polynôme} \Leftrightarrow P(a) = 0$$

Si $P(x) = (x-a)\,Q(x)$, il est évident que $P(a) = 0$. On admet la réciproque (si $P(a) = 0$, alors $P(x)$ est factorisable par $x - a$).

Voici quelques utilisations et conséquences de ce théorème :

• **Résolution d'équations ou d'inéquations de degré $\geqslant$ 3.** La mise en évidence d'une racine a par l'énoncé, ou l'existence d'une racine évidente a (le plus souvent 0, 1 ou -1), permet une mise en facteur par $x - a$, donc de « faire baisser le degré ».

> Soit $P(x) = x^3 + 5x^2 - 7x + 1$. Résoudre dans **R** :
>
> $$P(x) = 0 \ ; \quad P(x) \geqslant 0$$
>
> • $P(1) = 0$, donc P est factorisable par $x - 1$, donc
>
> $$P(x) = (x-1)\left(ax^2 + bx + c\right).$$
>
> Pour déterminer a, b, c, on peut
>
> – développer, puis procéder par identification ;
>
> – trouver les coefficients de proche en proche : $a = 1$, puis $b = 6$ en regroupant mentalement les deux termes en x^2 du développement...
>
> – utiliser la méthode de Horner, voir plus loin.
>
> On trouve
>
> $$P(x) = (x-1)\left(x^2 + 6x - 1\right),$$
>
> puis
>
> $$P(x) = 0 \Leftrightarrow x \in \left\{1, -3+\sqrt{10}, -3-\sqrt{10}\right\}$$
>
> • À l'aide d'un tableau de signes, on trouve
>
> $$P(x) \geqslant 0 \Leftrightarrow x \in \left[-3-\sqrt{10}, -3+\sqrt{10}\right] \cup [1, +\infty[$$

• **Méthode de Horner pour calculer $P(\alpha)$.** On considère

$$P(x) = a_n x^n + \cdots + a_1 x + a_0$$

Le polynôme $Q(x) = P(x) - P(\alpha)$ admet α pour racine, donc

$a_n x^n + \ldots a_1 x + a_0 - P(\alpha) = (x-\alpha)\left(b_{n-1}x^{n-1} + \cdots + b_1 x + b_0\right)$

$= b_{n-1}x^n + (b_{n-2} - \alpha b_{n-1})\,x^{n-1} + \cdots + (b_0 - \alpha b_1)\,x - \alpha b_0$

Par identification, on obtient le système d'équations

$$\left\{\begin{array}{c} b_{n-1} = a_n \\ b_{n-2} - \alpha b_{n-1} = a_{n-1} \\ \ldots \\ b_0 - \alpha b_1 = a_1 \\ -\alpha b_0 = a_0 - P(\alpha) \end{array}\right.$$

qui se résout en cascades : $b_{n-1} = a_n$, puis $b_{n-2} = a_{n-1} + \alpha b_{n-1}, \ldots$, $b_0 = a_1 + \alpha b_1$, $P(\alpha) = a_0 + \alpha b_0$. Disposition pratique (avec $n = 3$; les flèches indiquent une multiplication par α) :

a_3	a_2	a_1	a_0
	$+\alpha b_2$	$+\alpha b_1$	$+\alpha b_0$
$b_2 = a_3$	$= b_1$	$= b_0$	$= P(\alpha)$

La méthode de Horner se prête particulièrement bien à une programmation informatique et s'avère très économe en temps de calcul. Les variables d'entrée sont `n` (degré du polynôme, de type entier), `a` (suite des coefficients du polynôme par degrés croissants, de type tableau), `alpha`, de type réel. La variable de sortie est `Palpha`, de type réel.

```
 rogram horner ;
var n,i:integer ;al ha,Pal ha:real ;a:array[1..100]
of real ;
BEGIN
readln(n);for i:=0 to n do read(a[i]);readln(al ha);
Pal ha:=a[n];
for i:=n−1 downto 0 do
 Pal ha:=a[i]+al ha*Pal ha;
writeln(Pal ha);
END.
```

On compte avec cette méthode n additions et n multiplications, à comparer avec la programmation directe du calcul de $P(\alpha)$, qui conduirait à $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ multiplications et n additions.

• **Méthode de Horner pour factoriser par $x - \alpha$.** Dans le cas où α est une racine de P, la méthode de Horner continue de s'appliquer, elle aboutit au résultat 0, mais elle donne aussi les coefficients du polynôme $Q(x)$ tel que $P(x) = (x - \alpha)\,Q(x)$.

Exemple précédent : $P(x) = x^3 + 5x^2 - 7x + 1$, racine 1 :

1	5	-7	1
	$+1$	$+6$	$+(-1)$
1	$= 6$	$= -1$	$= 0$

D'où le résultat $P(x) = (x - 1)\left(x^2 + 6x - 1\right)$.

• On dit que α est une **racine d'ordre de multiplicité** n du polynôme P ssi $P(x) = (x - \alpha)^n Q(x)$, avec $Q(x)$ polynôme n'ayant pas α pour racine.

Une racine simple est une racine d'ordre de multiplicité 1, une racine double est une racine d'ordre de multiplicité 2. La somme des ordres de multiplicité des racines d'un polynôme de degré n est au plus égale à n.

• **Un polynôme de degré $n \geqslant 1$ admet au plus n racines**. Preuve par la contraposée, à savoir : si un polynôme admet plus de n racines, alors il n'est pas degré n. D'après le théorème de factorisation des polynômes, si le polynôme admet les racines $x_1, \ldots, x_{n+1}$, il est alors factorisable par $(x - x_1) \ldots (x - x_{n+1})$, et par conséquent il est de degré $> n$.

7. Manipulation des inégalités

7.1 Inégalités et opérations

• **Somme**

$$a \leqslant b \Rightarrow a + c \leqslant b + c\,; \qquad \left.\begin{array}{r} a \leqslant b \\ a' \leqslant b' \end{array}\right\} \Rightarrow a + a' \leqslant b + b'$$

En particulier, $a \leqslant b$ et $c \geqslant 0 \Rightarrow a \leqslant b + c$ et $a - c \leqslant b$.

Généralisation : $\forall i \in [\![1, n]\!],\ a_i \leqslant b_i \Rightarrow \displaystyle\sum_{i=1}^{n} a_i \leqslant \sum_{i=1}^{n} b_i$

Pour majorer une somme, on majore chaque terme de la somme.

• **Produit**

$$\left.\begin{array}{r} a \leqslant b \\ c \geqslant 0 \end{array}\right\} \Rightarrow ac \leqslant bc \quad ; \qquad \left.\begin{array}{r} a \leqslant b \\ c \leqslant 0 \end{array}\right\} \Rightarrow ac \geqslant bc$$

$$\left.\begin{array}{r} 0 \leqslant a \leqslant b \\ 0 \leqslant a' \leqslant b' \end{array}\right\} \Rightarrow 0 \leqslant aa' \leqslant bb'$$

On peut multiplier entre elles des inégalités de même sens portant sur des nombres **positifs**.

Ne pas oublier de renverser l'inégalité quand on multiplie par un nombre négatif.

Prudence si on ne connaît pas le signe du nombre par lequel on multiplie !

• **Opposé, inverse**

$$a \leqslant b \Rightarrow -a \geqslant -b \quad ; \quad 0 < a \leqslant b \Rightarrow \frac{1}{a} \geqslant \frac{1}{b}$$

Les inverses des nombres **positifs** se rangent dans l'ordre inverse.

• **Différence, quotient**

On ne peut rien dire de général, et on se gardera de soustraire ou diviser membre à membre des inégalités.

Pour encadrer une différence $x - y$, le mieux est d'encadrer $-y$, puis $x + (-y)$:

$$\left.\begin{array}{c} a \leqslant x \leqslant a' \\ b \leqslant y \leqslant b' \end{array}\right\} \Rightarrow \left.\begin{array}{c} a \leqslant x \leqslant a' \\ -b' \leqslant -y \leqslant -b \end{array}\right\} \Rightarrow a - b' \leqslant x - y \leqslant a' - b$$

Pour encadrer un quotient de nombres **positifs** :

$$\left.\begin{array}{c} 0 < a \leqslant x \leqslant a' \\ 0 < b \leqslant y \leqslant b' \end{array}\right\} \Rightarrow \left.\begin{array}{c} 0 < a \leqslant x \leqslant a' \\ 0 < \dfrac{1}{b'} \leqslant \dfrac{1}{y} \leqslant \dfrac{1}{b} \end{array}\right\} \Rightarrow 0 < \frac{a}{b'} \leqslant \frac{x}{y} \leqslant \frac{a'}{b}$$

Pour majorer un quotient de nombres positifs, on majore le numérateur et on minore le dénominateur.

7.2 Inégalités et fonctions

Principe général.
Soit f définie sur l'intervalle I, et $a, b \in I$. Alors
Si f est croissante : $a \leqslant b \Rightarrow f(a) \leqslant f(b)$
Si f est décroissante : $a \leqslant b \Rightarrow f(a) \geqslant f(b)$
Si f est continue et strictement croissante : $a \leqslant b \Leftrightarrow f(a) \leqslant f(b)$
Si f est continue et strictement décroissante : $a \leqslant b \Leftrightarrow f(a) \geqslant f(b)$

Dans les deux derniers cas, il y a équivalence car f est une bijection, et f^{-1} a même sens de variation que f.

Une fonction croissante conserve le sens des inégalités, une fonction décroissante le renverse.

$0 \leqslant a \leqslant b \Rightarrow a^2 \leqslant b^2$

$0 \leqslant a \leqslant b \Leftrightarrow \sqrt{a} \leqslant \sqrt{b}$

$0 \leqslant a \leqslant b \Rightarrow a^r \leqslant b^r$ pour tout nombre réel r positif

$0 < a \leqslant b \Leftrightarrow \ln a \leqslant \ln b$; $a \leqslant b \Leftrightarrow e^a \leqslant e^b$

$0 < a < b \Leftrightarrow \dfrac{1}{a} > \dfrac{1}{b} > 0$

• Attention à la fonction $x \mapsto x^2$:

– $0 \leqslant a \leqslant b \Rightarrow a^2 \leqslant b^2$, mais $a^2 \leqslant b^2 \Rightarrow \sqrt{a^2} \leqslant \sqrt{b^2}$, c'est-à-dire seulement : $a^2 \leqslant b^2 \Rightarrow |a| \leqslant |b|$

– Si $a \geqslant 0$, $b \geqslant 0$, alors $a^2 \leqslant b^2 \Rightarrow a \leqslant b$ (et il y a équivalence).

– On notera en particulier, avec $b \geqslant 0$:

$$x^2 \leqslant b \Leftrightarrow -\sqrt{b} \leqslant x \leqslant \sqrt{b}\,; \quad x^2 \geqslant b \Leftrightarrow x \leqslant -\sqrt{b} \text{ ou } x \geqslant \sqrt{b}$$

• À noter également, avec n entier $\geqslant 2$:

$$0 \leqslant x \leqslant 1 \Rightarrow 0 \leqslant x^n \leqslant x \leqslant \sqrt{x} \leqslant 1$$

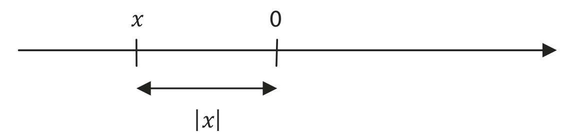
$$x \geqslant 1 \Rightarrow x^n \geqslant x \geqslant \sqrt{x} \geqslant 1$$

Inégalités et valeur absolue

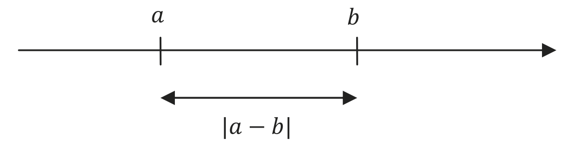
• $|x| = \text{Max}\,\{x, -x\}$ (le plus grand des deux nombres $x, -x$).

• $|x| = \begin{cases} x & \text{si} \quad x \geqslant 0 \\ -x & \text{si} \quad x \leqslant 0 \end{cases}$

• $|x|$ est la distance du point x au point 0 de la droite réelle :



• $|a - b|$ est la distance du point a au point b :



• Pour tout $x \in \mathbf{R}$: $|x| \geqslant 0$, et $|x| = 0 \Leftrightarrow x = 0$.

• Pour tout a, b dans $\mathbf{R}$: $|a + b| \leqslant |a| + |b|$ (**inégalité triangulaire**).

Généralisation :

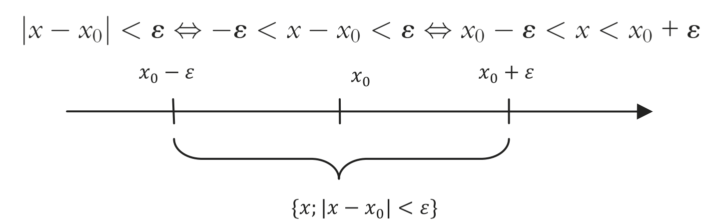
$$\left|\sum_{i=1}^{n} a_i\right| \leqslant \sum_{i=1}^{n} |a_i|$$

Attention, on a seulement $|a - b| \leqslant |a| + |b|$.

• $\lvert A\rvert \leqslant B \Leftrightarrow -B \leqslant A \leqslant B$; $\lvert A\rvert \geqslant B \Leftrightarrow A \geqslant B$ ou $A \leqslant -B$

Ainsi, avec $\varepsilon > 0$:

$$\lvert x - x_0\rvert < \varepsilon \Leftrightarrow -\varepsilon < x - x_0 < \varepsilon \Leftrightarrow x_0 - \varepsilon < x < x_0 + \varepsilon$$



Partie 1

Analyse

Étude de fonctions

1

Vocabulaire de base

Soit f une fonction numérique de la variable réelle, c'est-à-dire une fonction de **R** dans **R**.

Soit D une partie non vide de **R**.

On dit que f est **définie sur** D ssi f est une application de D dans **R**.

Soit f une application définie sur D.

On dit que f est

paire ssi $\forall x \in D, \quad -x \in D \quad \text{et} \quad f(-x) = f(x)$,

impaire ssi $\forall x \in D, \quad -x \in D \quad \text{et} \quad f(-x) = -f(x)$.

Soit I un intervalle non vide inclus dans D.

On dit que f est

• **croissante**, resp. **décroissante** sur I ssi, pour tout a, b dans I :

$$a \leqslant b \Rightarrow f(a) \leqslant f(b), \quad \text{resp.} \quad a \leqslant b \Rightarrow f(a) \geqslant f(b)$$

• **strictement croissante**, resp. **strictement décroissante** sur I ssi, pour tout a, b dans I :

$$a < b \Rightarrow f(a) < f(b), \quad \text{resp.} \quad a < b \Rightarrow f(a) > f(b)$$

• **monotone** sur I ssi f est croissante ou décroissante sur I ;

• **strictement monotone** sur I ssi f est strictement croissante ou strictement décroissante sur I ;

• **majorée par** M, resp. **minorée par** m sur I ssi

$$\forall x \in I, \ f(x) \leqslant M \quad \text{resp.} \quad \forall x \in I, \ f(x) \geqslant m$$

M est alors un **majorant** et m est un **minorant** de f sur I.

• **bornée** sur I ssi f est majorée et minorée sur I.

Les paragraphes 1 à 5 concernent les fonctions de $\mathbf{R}$ *dans* $\mathbf{R}$.

Le paragraphe 6 concerne les fonctions de deux variables (fonctions de $\mathbf{R}^2$ *dans* $\mathbf{R}$).

1. Recherche de limites

1.1 Définitions

Soit I un intervalle de $\mathbf{R}$, $x_0 \in I$, f une fonction définie sur D, avec $D = I$ ou $D = I \setminus \{x_0\}$, et ℓ un nombre réel. On pose

- $\lim\limits_{x_0} f = \ell$ ssi :

$$\forall \varepsilon > 0,\ \exists \alpha > 0, \quad x \in D \text{ et } |x - x_0| < \alpha \Rightarrow |f(x) - \ell| < \varepsilon.$$

- $\lim\limits_{x_0^+} f = \ell$ ssi il existe $x \in I$ tel que $x > x_0$, et :

$$\forall \varepsilon > 0,\ \exists \alpha > 0, \quad 0 < x - x_0 < \alpha \Rightarrow |f(x) - \ell| < \varepsilon.$$

- $\lim\limits_{x_0} f = +\infty$ ssi :

$$\forall A \in \mathbf{R},\ \exists \alpha > 0, \quad x \in D \text{ et } |x - x_0| < \alpha \Rightarrow f(x) > A$$

On définit de façon analogue la limite à gauche en $x_0 \left(\lim\limits_{x_0^-} f\right)$, la limite en $+\infty$, en $-\infty$, ces limites étant éventuellement $+\infty$, $-\infty$.

On a aussi les notations :

$$\lim_{x \to x_0} f(x) = \ell \ ; \qquad \lim_{x \to x_0^+} f(x) = \ell \ ; \qquad \lim_{x \to x_0, x > x_0} f(x) = \ell \ldots$$

La notation $f(x) \underset{x \to x_0}{\longrightarrow} \ell$ s'avère très pratique, mais il faut veiller à ne pas séparer les deux flèches, qui n'ont de signification que conjointement.

1.2 Opérations sur les limites

Limites usuelles

On rappelle les limites suivantes :

$$\text{Avec } r > 0 : \lim_{x \to +\infty} x^r = +\infty \ ; \qquad \lim_{x \to +\infty} \frac{1}{x^r} = 0$$

$$\text{Avec } n \in \mathbf{N}^*, \lim_{x \to -\infty} x^n = \begin{cases} +\infty & \text{si } n \text{ est pair} \\ -\infty & \text{si } n \text{ est impair} \end{cases}$$

$$\lim_{x \to 0} \ln x = -\infty \ ; \quad \lim_{x \to +\infty} \ln x = +\infty \ ; \quad \lim_{x \to -\infty} e^x = 0 \ ; \quad \lim_{x \to +\infty} e^x = +\infty$$

Dans la suite du paragraphe, x_0, ℓ, ℓ' sont des nombres réels. b, b', b'' sont mis à la place d'un des symboles $x_0, x_0^+, x_0^-, +\infty, -\infty$.

Limite d'une somme, d'un produit, d'un quotient

Théorème. Soit f et g telles que

$$\lim_b f = \ell, \quad \lim_b g = \ell'$$

Alors :

$$\lim_b (f + g) = \ell + \ell'$$

$$\lim_b (fg) = \ell\ell'$$

$$\lim_b \left(\frac{f}{g}\right) = \frac{\ell}{\ell'} \text{ si de plus } \ell' \neq 0$$

Si une des limites est infinie ou si $\ell' = 0$, on a des résultats partiels.
RS signifie qu'on utilise la règle des signes, FI signale une forme indéterminée :

$\lim\limits_b u$	$\lim\limits_b v$	$\lim\limits_b (u+v)$	$\lim\limits_b (uv)$	$\lim\limits_b \left(\frac{u}{v}\right)$
$\ell \neq 0$	0	ℓ	0	$\pm\infty$ RS
0	0	0	0	FI
$\ell \neq 0$	$\pm\infty$	$\pm\infty$	$\pm\infty$ RS	0
0	$\pm\infty$	$\pm\infty$	FI	0
$\pm\infty$	$\ell' \neq 0$	$\pm\infty$	$\pm\infty$ RS	$\pm\infty$ RS
$\pm\infty$	0	$\pm\infty$	FI	0
$+\infty$	$-\infty$	FI	$-\infty$	FI
$+\infty$	$+\infty$	$+\infty$	$+\infty$	FI

Limite d'une fonction composée

$$\lim_{x \to b} f(x) = b' \;;\; \lim_{y \to b'} g(y) = b'' \Rightarrow \lim_{x \to b} g \circ f(x) = b''$$

- $\lim\limits_{x \to 0^+} e^{\frac{1}{x}} = +\infty$ car $\lim\limits_{x \to 0^+} \frac{1}{x} = +\infty$ et $\lim\limits_{y \to +\infty} e^y = +\infty$
- $\lim\limits_{x \to 0^-} e^{\frac{1}{x}} = 0$ car $\lim\limits_{x \to 0^-} \frac{1}{x} = -\infty$ et $\lim\limits_{y \to -\infty} e^y = 0$

Passage à la limite dans les inégalités

$$\text{Si } f \leqslant g,\quad \lim_b f = \ell,\quad \lim_b g = \ell',\quad \text{alors } \ell \leqslant \ell'$$

$$\text{Si } f \leqslant g \leqslant h,\quad \lim_b f = \lim_b h = \ell,\quad \text{alors } \lim_b g = \ell$$

$$\text{Si } |f(x) - \ell| \leqslant g(x) \text{ et } \lim_b g = 0,\quad \text{alors } \lim_b f = \ell$$

Les deux premières formules restent valables si ℓ, ℓ' appartiennent à $\{-\infty, +\infty\}$. En particulier :

$$\text{Si } f \leqslant g \text{ et } \lim_b f = +\infty,\quad \text{alors } \lim_b g = +\infty$$

$$\text{Si } f \leqslant g \text{ et } \lim_b g = -\infty,\quad \text{alors } \lim_b f = -\infty$$

Notons ce **théorème d'existence** :

Soit f une fonction croissante (resp. décroissante) et majorée par M (resp. minorée par m) sur l'intervalle $[a, b[$, avec $a < b \leqslant +\infty$. Alors f admet une limite finie ℓ en b, et $\ell \leqslant M$ (resp. $\ell \geqslant m$).

Ce théorème est à rapprocher du théorème analogue sur les suites numériques, voir § 2.1.2. Il ne permet pas de déterminer la limite ℓ.

1.3 Négligeabilité

Dans la suite du chapitre, on parlera de propriétés vérifiées **au voisinage de** b, c'est-à-dire sur un ensemble non vide du type $[a, x_0[\cap]x_0, a']$ si $b = x_0$, un intervalle (non vide) $[a, +\infty[$ si $b = +\infty$, ou $]-\infty\,; a]$ si $b = -\infty$

Définition. On dit que f est **négligeable devant** g en b , et on note $f \underset{b}{=} \circ(g)$, ssi

$$\lim_{x \to b} \frac{f(x)}{g(x)} = 0$$

La définition adoptée suppose que g ne s'annule pas au voisinage de x_0. Cela ne pose pas de difficultés dans la pratique.

Négligeabilités classiques. Pour $\alpha > 0$:

$$\lim_{x\to+\infty} \frac{\ln x}{x^\alpha} = 0 \qquad \left(\text{et donc } \ln x \underset{+\infty}{=} \circ\left(x^\alpha\right)\right)$$

$$\lim_{x\to+\infty} \frac{e^x}{x^\alpha} = +\infty \qquad \left(\text{et donc } x^\alpha \underset{+\infty}{=} \circ\left(e^x\right)\right)$$

$$\lim_{x\to0} \left(x^\alpha \ln x\right) = 0 \qquad \left(\text{et donc } \ln x \underset{0}{=} \circ\left(\frac{1}{x^\alpha}\right)\right)$$

Mémorisez soigneusement ces limites, elles sont d'usage constant. Elles n'ont pas à être justifiées, elles font partie des connaissances de base. Au besoin, vous évoquerez les « négligeabilités classiques ».

Vous pouvez retenir aussi, pour $n \in \mathbf{N}$: $\lim\limits_{x\to-\infty} x^n e^x = 0$.

On a aussi : $\lim\limits_{x\to+\infty} \left(\dfrac{x^\alpha}{e^x}\right) = 0 \ldots$

1.4 Équivalence

Définition. On dit que f est **équivalente** à g en b, et on note $f \underset{b}{\sim} g$, ssi

$$\lim_{x\to b} \frac{f(x)}{g(x)} = 1$$

La définition adoptée suppose que g ne s'annule pas au voisinage de x_0. Cela ne pose pas de difficultés dans la pratique.

Équivalents classiques.

$$\ln x \underset{1}{\sim} x - 1 \;; \qquad \ln(1+h) \underset{0}{\sim} h \;; \qquad e^x - 1 \underset{0}{\sim} x$$

Considérer $\ln'(1) = 1$ pour montrer la première équivalence (voir la définition de $f'(x_0)$ §1.3.1). Poser alors $x = 1 + h$ pour montrer la deuxième. La troisième équivalence se démontre en considérant $\exp'(0) = 1$.

Propriétés

- $u \underset{b}{\sim} v \;\Leftrightarrow\; u = v + \circ_b(v) \;\Leftrightarrow\; u = v(1+\varepsilon)$ avec $\lim\limits_{b} \varepsilon = 0$
- $\left.\begin{array}{l} u_1 \sim v_1 \\ u_2 \sim v_2 \end{array}\right\} \Rightarrow u_1 u_2 \sim v_1 v_2 \;;\; \dfrac{u_1}{u_2} \sim \dfrac{v_1}{v_2}$

- Pour tout $\alpha \in \mathbf{R}$ tel que u^α et v^α existent : $u \sim v \Rightarrow u^\alpha \sim v^\alpha$
- $u \sim v \Rightarrow v \sim u$; $u \sim v$ et $v \sim w \Rightarrow u \sim w$

Pour établir la première propriété, posez $\frac{u}{v} = 1 + \varepsilon$. Les propriétés suivantes sont des conséquences directes de la définition.

La première propriété permet de « voir » immédiatement des équivalents :

- $x - \ln x \underset{+\infty}{\sim} x$ car $\ln x \underset{+\infty}{=} \circ(x)$; $e^x - x^2 \underset{x \to +\infty}{\sim} e^x$ car $x^2 \underset{+\infty}{=} \circ(e^x)$
- Un polynôme est équivalent en $\pm\infty$ à son terme de plus haut degré.

Les deux propriétés suivantes signifient que les équivalences « passent » dans les produits, les quotients, les élévations à la puissance. Mais attention, vous ne devez pas croire que c'est le cas pour la somme, le logarithme, l'exponentielle :

- En 0, $x^2 + x^3 \sim x^2$; $-x^2 \sim x^2 + x^4$; mais $x^3 \nsim x^4$.
- En 0, $1 + x^2 \sim 1 + x$, mais $\ln\left(1 + x^2\right) \nsim \ln(1 + x)$, car $\ln\left(1 + x^2\right) \sim x^2, \ln(1 + x) \sim x$, et, d'après la dernière propriété, si on avait $\ln\left(1 + x^2\right) \sim \ln(1 + x)$, on aurait $x^2 \sim x$.
- En $+\infty$, $x^2 + x \sim x^2$, mais $e^{x^2} \nsim e^x$ (faire le quotient).

Utilisation

$$\left.\begin{array}{l} u \underset{b}{\sim} v \\ \lim\limits_{b} v = \ell \end{array}\right\} \Rightarrow \lim_{b} u = \ell \;;\; u \underset{b}{\sim} \ell \in \mathbf{R}^* \Rightarrow \lim_{b} u = \ell$$

Pour déterminer la limite d'une fonction en un point, on peut essayer de lui trouver un équivalent plus simple, dont on cherche la limite.

- Avec $f(x) = \dfrac{x^3}{e^x - 1}$, $\lim\limits_{0} f = \lim\limits_{+\infty} f = 0$.

En effet $\dfrac{x^3}{e^x - 1} \underset{0}{\sim} \dfrac{x^3}{x} = x^2 \underset{x \to 0}{\longrightarrow} 0$, donc $\lim\limits_{0} f = 0$;

$\dfrac{x^3}{e^x - 1} \underset{+\infty}{\sim} \dfrac{x^3}{e^x} \underset{x \to +\infty}{\longrightarrow} 0$ (négligeabilité classique), donc $\lim\limits_{+\infty} f = 0$.

- $\lim_{x\to-\infty} e^{-x}\ln(1+e^x) = 1$ car $e^{-x}\ln(1+e^x) \underset{-\infty}{\sim} e^{-x}e^x = 1$.

On a utilisé $\lim_{x\to-\infty} e^x = 0$ et $\ln(1+y) \underset{0}{\sim} y$.

Attention à utiliser les équivalents à bon escient :

- $\lim_{x\to 0} \dfrac{-x\ln x}{1+x^2} = 0$ car $\lim_{x\to 0} x\ln x = 0$: les équivalents sont ici inutiles.
- $\lim_{x\to+\infty} \dfrac{-x\ln x}{1+x^2} = 0$ car $\dfrac{-x\ln x}{1+x^2} \underset{+\infty}{\sim} \dfrac{-x\ln x}{x^2} = -\dfrac{\ln x}{x} \underset{x\to 0}{\longrightarrow} 0$: ne pas chercher d'équivalent à ln en $+\infty$.
- $\lim_{x\to+\infty} \dfrac{\ln(1+x)}{x} = 0$: ne pas utiliser $1+x \underset{+\infty}{\sim} x$: les équivalents ne « passent » pas aux logarithmes. Le mieux est d'écrire :

$$\frac{\ln(1+x)}{x} = \frac{\ln\left[x\left(1+\frac{1}{x}\right)\right]}{x} = \frac{\ln x + \ln\left(1+\frac{1}{x}\right)}{x}$$
$$= \frac{\ln x}{x} + \frac{\ln\left(1+\frac{1}{x}\right)}{x}$$

ce qui permet de conclure.

- $u \underset{b}{\sim} v$ et $\lim_b v = \ell \Rightarrow \lim_b u = \ell$, mais deux fonctions ayant même limite ne sont pas toujours équivalentes ! Par exemple :

$$\lim_{x\to 0} 2x^2 = \lim_{x\to 0} x^2 = 0, \quad \text{mais} \quad 2x^2 \underset{0}{\nsim} x^2.$$

- $u \underset{b}{\sim} \ell \in \mathbf{R}^* \Rightarrow \lim_b u = \ell$: attention, l'écriture $u \underset{b}{\sim} 0$ n'a aucune signification, et est à proscrire absolument ! (Idem pour $u \underset{b}{\sim} +\infty$)
- Réciproquement, $\lim_b u = \ell \in \mathbf{R}^* \Rightarrow u \underset{b}{\sim} \ell$.

1.5 Utilisation des développements limités

Développements limités usuels

À partir de l'inégalité de Taylor-Lagrange (§ 3.2.5), on trouve et on retient les **développements limités** (DL) **d'ordre** n **au voisinage de 0** suivants :

ε est une fonction de limite nulle en 0, $n \in \mathbf{N}^*$, $\alpha \in \mathbf{R}$.

$$e^x = 1 + x + \tfrac{x^2}{2!} + \cdots + \frac{x^n}{n!} + x^n\varepsilon(x)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \cdots + (-1)^{n+1}\frac{x^n}{n} + x^n\varepsilon(x)$$

$$(1+x)^\alpha = 1 + \alpha x + \cdots + \frac{\alpha(\alpha-1)(\alpha-n+1)}{n!}x^n + x^n\varepsilon(x)$$

• Les DL les plus fréquemment utilisés sont les suivants :

$$e^x = 1 + x + \frac{x^2}{2!} + x^2\varepsilon(x) \;; \quad e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + x^3\varepsilon(x)$$

$$\ln(1+x) = x - \frac{x^2}{2} + x^2\varepsilon(x) \;; \quad \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + x^3\varepsilon(x)$$

• À noter le DL : $\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + x^n\varepsilon(x)$, qui se retrouve à partir de l'identité géométrique.

• On peut écrire, de façon équivalente, $e^x = 1 + x + \frac{x^2}{2} + \circ\left(x^2\right)$, par exemple. Mais l'écriture adoptée ici semble plus maniable.

Pour la mémorisation, le troisième DL encadré est à rapprocher de la formule du binôme (le coefficient de x^n comporte n facteurs en descendant à partir de α, resp. de n, pour le numérateur, resp. le dénominateur).

On obtient un DL au voisinage de $x_0 \in \mathbf{R}$ en posant $x = x_0 + h$, et au voisinage de $\pm\infty$ en posant $x = \frac{1}{h}$.

• Ainsi, avec $x = 1 + h$, on a

$$\begin{aligned} x\ln x &= (1+h)\ln(1+h) \\ &= (1+h)\left(h - \frac{h^2}{2} + h^2\varepsilon(h)\right) = h - \frac{h^2}{2} + h^2\varepsilon(h) \\ &= x - 1 + \frac{(x-1)^2}{2} + (x-1)^2\varepsilon(x-1) \text{ (DL en 1)} \end{aligned}$$

• Et avec $x = \dfrac{1}{h}$:

$$\begin{aligned} xe^{-\frac{1}{x}} &= \frac{1}{h}\left(1 - h + \frac{h^2}{2} + h^2\varepsilon(h)\right) = \frac{1}{h} - 1 + \frac{h}{2} + h\varepsilon(h) \\ &= x - 1 + \frac{1}{2x} + \frac{1}{x}\varepsilon\left(\frac{1}{x}\right) \text{ (« DL » en } +\infty) \end{aligned}$$

Propriétés

• Soit le DL d'ordre n de f en 0 : $f(x) = P_n(x) + x^n\varepsilon(x)$, avec P_n polynôme de degré $\leqslant n$. Alors, **si** $\lim_0 u = 0$, on a :

$$f(u(x)) = P_n(u(x)) + (u(x))^n\varepsilon(x)$$

On devrait écrire $\varepsilon(u(x))$, mais cela n'apporte aucune information, car la seule chose que l'on sait de ε, c'est que c'est une fonction de limite nulle en 0.

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} + x^3\varepsilon(x) \quad ; \quad e^{x^2} = 1 + \frac{x^2}{2} + \frac{x^4}{6} + x^4\varepsilon(x)$$

• Dans le développement du produit de deux DL, on n'écrit pas les termes qui sont « absorbés » par le terme en $x^n\varepsilon(x)$ dont l'ordre est le plus petit.

$$\frac{e^x}{1-x} = e^x \times \frac{1}{1-x} = \left(1 + \frac{x^2}{2} + x^2\varepsilon(x)\right)\left(1 + x + x^2 + x^2\varepsilon(x)\right)$$

$$= 1 + x + x^2 + x^2\varepsilon(x) + \frac{x^2}{2}$$

Il est inutile d'écrire les autres termes $\left(\frac{x^3}{2}, \frac{x^4}{2}, \ldots\right)$ qui sont tous de la forme $x^2\varepsilon(x)$ et n'apportent donc aucune précision supplémentaire. Il reste alors à achever les calculs.

Utilisation

On utilise les DL dans la recherche des limites quand les équivalents s'avèrent inopérants.

Prouvons que $\lim_0 f = \frac{1}{2}$, avec, pour $x \in \mathbf{R}^*$:

$$f(x) = \frac{e^x - 1 - x}{x^2}$$

Il s'agit d'une forme indéterminée ; on ne peut utiliser l'équivalent $e^x - 1 \underset{0}{\sim} x$ car ça ne « passe » pas dans les sommes : le numérateur serait équivalent à 0 ? Mais on peut remplacer e^x par son DL, l'ordre 2 suffit :

$$f(x) = \frac{1 + x + \frac{x^2}{2} + x^2\varepsilon(x) - 1 - x}{x^2} = \frac{\frac{x^2}{2} + x^2\varepsilon(x)}{x^2} = \frac{1}{2} + \varepsilon(x)$$

Ce qui permet de conclure.

Voir d'autres utilisations, § 2.3.2 (théorème de prolongement de la dérivée), § 2.4.4 (étude des branches infinies).

1.6 Autres techniques

• Limite de $f(x) = \ln\left(e^x - e^{-x}\right) - x$ en $+\infty$. Il y a forme indéterminée, on ne peut utiliser d'équivalents (qui ne passent ni dans les logarithmes ni dans les sommes), et un DL ne semble pas très naturel. L'idée est d'utiliser $x = e^{\ln x}$, puis les règles de calcul pour les logarithmes :

$$f(x) = \ln\left(e^x - e^{-x}\right) - \ln e^x = \ln \frac{e^x - e^{-x}}{e^x} = \ln\left(1 - e^{-2x}\right)$$

tend vers 0 quand x tend vers $+\infty$.

• Limite de $f(x) = \sqrt{4 + (x+1)^2} - (x+1)$ en $+\infty$. L'utilisation d'un DL est possible, mais la technique de la « quantité conjuguée » est plus rapide :

$$\begin{aligned} f(x) &= \frac{\left(\sqrt{4 + (x+1)^2} - (x+1)\right)\left(\sqrt{4 + (x+1)^2} + (x+1)\right)}{\sqrt{4 + (x+1)^2} + (x+1)} \\ &= \frac{4 + (x+1)^2 - (x+1)^2}{\sqrt{4 + (x+1)^2} + (x+1)} = \frac{4}{\sqrt{4 + (x+1)^2} + (x+1)} \end{aligned}$$

tend vers 0 quand x tend vers $+\infty$.

• Limites en $+\infty$ et $-\infty$ de $f(x) = \frac{x+1}{\sqrt{x^2+1}}$. L'utilisation des équivalents est possible. On peut aussi mettre le terme dominant en facteur à l'intérieur du radical :

$$\frac{x+1}{\sqrt{x^2+1}} = \frac{x+1}{\sqrt{x^2\left(1 + \frac{1}{x^2}\right)}} = \frac{x+1}{|x|\sqrt{\left(1 + \frac{1}{x^2}\right)}}$$

ce qui permet de conclure en remplaçant $|x|$ par x ou $-x$, puis en utilisant les équivalents.

1.7 Branches infinies, éléments graphiques

Asymptotes

• Si $\lim_{x_0} f = \pm\infty$ ou $\lim_{x_0+} f = \pm\infty$ ou $\lim_{x_0-} f = \pm\infty$, la droite d'équation $x = x_0$ est asymptote à la courbe (asymptote verticale) ($x_0 \in \mathbf{R}$).

• Si $\lim_{\pm\infty} f = b \in \mathbf{R}$, la droite d'équation $y = b$ est asymptote à la courbe (asymptote horizontale).

• Si $f(x) = ax + b + \varepsilon(x)$ et $\lim_{\pm\infty} \varepsilon = 0$, la droite d'équation $y = ax + b$ est asymptote à la courbe (oblique si $a \neq 0$).

Étude systématique dans le cas $\lim_{\pm\infty} f = \pm\infty$

• Si $\lim_{x\to\pm\infty} \frac{f(x)}{x} = 0$:

C_f admet une branche parabolique de direction (Ox)

• Si $\lim_{x\to\pm\infty} \frac{f(x)}{x} = \pm\infty$:

C_f admet une branche parabolique de direction (Oy)

• Si $\lim_{x\to\pm\infty} \frac{f(x)}{x} = a \in \mathbf{R}^*$:

– dans le cas général : C_f a pour direction asymptotique la droite d'équation $y = ax$

– si $\lim_{x\to\pm\infty} (f(x) - ax) = b \in \mathbf{R}$:

C_f admet la droite d'équation $y = ax + b$ pour asymptote

– si $\lim_{x\to\pm\infty} (f(x) - ax) = \pm\infty$:

C_f admet une branche parabolique de direction la droite d'équation $y = ax$

Avant de faire l'étude systématique ci-dessus, s'assurer de n'être pas dans le cas $f(x) = ax + b + \varepsilon(x)$ évoqué plus haut. L'énoncé peut demander de déterminer a, b et ε par identification.

On peut être amené à utiliser les DL pour l'étude d'une branche infinie.

• Pour $x \neq -1, f(x) = \frac{x^2}{x-1} = ax + b + \frac{c}{x-1} = \frac{ax^2+(b-a)x+c-b}{x-1} \Leftrightarrow$ $a = 1, b - a = 0, c - b = 0$; donc $f(x) = x + 1 + \frac{1}{x-1}$

On a donc une asymptote D d'équation $y = x+1$, et C_f est au-dessus de D pour les points d'abscisse > -1, en-dessous pour les points d'abscisse < -1.

• Avec $x = \frac{1}{h}$, au voisinage de $+\infty$:

$$f(x) = xe^{-\frac{1}{x}} = \frac{1}{h}\left(1 - h + \frac{h^2}{2} + h^2\varepsilon(h)\right) = \frac{1}{h} - 1 + \frac{h}{2} + h\varepsilon(h)$$

$$= x - 1 + \frac{1}{2x} + \frac{1}{x}\varepsilon\left(\frac{1}{x}\right).$$

D d'équation $y = x - 1$ est asymptote à C_f.

Au voisinage de $+\infty$, C_f est au-dessus de D car $\frac{1}{x}\varepsilon\left(\frac{1}{x}\right)$ est négligeable devant $\frac{1}{2x}$, qui donne donc le signe de $f(x) - (x-1)$ pour les grandes valeurs de x.

Autres éléments graphiques

$f : D \to \mathbf{R}$ est **paire** ssi $\forall x \in D, -x \in D$ et $f(-x) = f(x)$.

On étudie f sur $D \cap \mathbf{R}^+$. Dans un repère orthogonal (xOy), C_f est symétrique par rapport à (Oy).

$f : D \to \mathbf{R}$ est **impaire** ssi $\forall x \in D, -x \in D$ et $f(-x) = -f(x)$.

On étudie f sur $D \cap \mathbf{R}^+$. Dans un repère orthogonal (xOy), C_f est symétrique par rapport à O.

Pour le tracé de C_f, commencez (s'il y a lieu) par placer les asymptotes, et les points à tangentes remarquables, c'est-à-dire : points à tangente horizontale, points anguleux (où les dérivées à gauche et à droite sont différentes), voir § 1.3.2 ; points d'inflexion si ceux-ci sont demandés, voir § 1.3.3. Le point d'abscisse 0 de C_f (s'il existe) est souvent intéressant à placer.

2. Continuité

2.1 Définitions

Soit I un intervalle de $\mathbf{R}$, et soit $x_0 \in I$.

• Soit f une fonction définie sur I. On dit que f est **continue en** x_0 ssi

$$\lim_{x_0} f = f(x_0)$$

On définit de même la continuité à droite et à gauche en x_0.

f est continue en $x_0 \in]a\,;b[$ ssi f continue à droite et à gauche en x_0.

• On dit que f est **continue sur** I ssi f est continue en tout point de I.

f est continue sur $[a\,;b]$ ssi f est continue sur $]a\,;b[$, continue à droite en b et continue à gauche en a.

• Soit f un fonction définie sur $I \setminus \{x_0\}$. On dit que f est **prolongeable par continuité en** x_0 ssi f admet une limite finie en x_0.

En posant $f(x_0) = \lim\limits_{x_0} f$, on obtient un prolongement de f à I, et ce prolongement est continu en x_0.

2.2 Opérations sur les fonctions continues

Théorème

• Les fonctions polynômes, la fonction exponentielle sont continues sur **R**. La fonction ln est continue sur $]0\,;\,+\infty[$.

La fonction racine carrée est continue sur $[0\,;\,+\infty[$, ainsi que les fonctions $x \mapsto x^r$, avec $r > 0$.

• La somme, le produit, le quotient avec le dénominateur qui ne s'annule pas de deux fonctions continues sur I sont des fonctions continues sur I.

• Si u est continue sur I et f continue sur $u(I)$, alors la composée $f \circ u$ est continue sur I.

On utilise ce théorème pour montrer qu'une fonction est continue sur un intervalle. Il s'étend sans difficulté au cas où f est définie sur une réunion d'intervalles ($\mathbf{R}^*$ par exemple). Mais on peut être amené à revenir à la définition si la fonction f est définie de façon particulière en un (ou plusieurs) points, ou s'il y a problème de « raccordement » :

• Définition particulière en un point : soit f la fonction définie sur $[0\,;\,+\infty[$ par

$$f(x) = \frac{x\ln x}{x-1} \text{ si } x \in\,]0\,;1[\,\cup\,]1\,;\,+\infty[\quad ; \quad f(0)=0 \quad ; \quad f(1)=1$$

– f est continue sur $]0\,;1[\,\cup\,]1\,;\,+\infty[$ en tant que quotient de deux fonctions continues avec le dénominateur qui ne s'annule pas.

– f est continue en 0, car $\lim\limits_{x\to 0} x\ln x = 0$, donc

$$\lim_0 f = 0 = f(0)\,.$$

– f est continue en 1, car $\ln x \underset{1}{\sim} x-1$, donc $f(x) \underset{1}{\sim} x \underset{x\to 1}{\longrightarrow} 1 = f(1)$.

– Donc f est continue sur $[0\,;\,+\infty[$.

• Problème de raccordement : soit f la fonction définie sur **R** par

$$f(x) = x(1-x) \text{ si } x \in [0\,;1] \;;\quad f(x) = 0 \text{ sinon.}$$

– f est continue sur $[0\,;1]$ (fonction polynôme) et sur $]-\infty\,;0[$ et $]1\,;\,+\infty[$ (fonction polynôme).

– f est continue en 0 car $\lim\limits_{0^-} f = \lim\limits_{0^+} f = 0 = f(0)$.

– f est continue en 1 car $\lim\limits_{1^-} f = \lim\limits_{1^+} f = 0 = f(1)$.

– Donc f est continue sur **R**.

Vous utiliserez ce théorème de façon très précise, mais brève :

• Repérer si la fonction est une somme, un produit, une composée... de fonctions continues et la traiter en tant que telle.

• Si la fonction est un quotient, ne pas oublier de mentionner – et de vérifier – que le dénominateur ne s'annule pas.

• Si la fonction est une composée $f \circ u$, vérifier que $u(I)$ est inclus dans un ensemble sur lequel f est continue.

Exemple : $x \mapsto x^2 + 1$ est continue **et** (strictement) **positive** sur **R**, ln est continue sur $]0\,;\,+\infty[$, donc la composée $x \mapsto \ln\left(x^2+1\right)$ est continue sur **R**.

• Ne pas évoquer vaguement « somme, produit et quotient » de fonctions continues, ou « composée » de fonctions continues (« composée » a ici un sens précis), ou encore les « fonctions usuelles », qui ne sont pas toutes continues partout !

• Enfin, ne pas entrer trop dans les détails, sous peine d'avoir une rédaction interminable où le risque d'erreur augmente. L'essentiel est que vous ayez compris comment on « fabrique » une fonction continue à partir de fonctions continues plus simples.

2.3 Propriétés des fonctions continues

Théorème des valeurs intermédiaires

L'image d'un intervalle par une fonction continue est un intervalle.

Entre deux valeurs que prend une fonction continue sur un intervalle, la fonction prend toutes les valeurs intermédiaires. Théorème admis.

Corollaire : théorème de la bijection monotone

Soit f une fonction **continue** et **strictement monotone** sur l'**intervalle** I. Alors f réalise une bijection de l'intervalle I sur l'intervalle $J = f(I)$. De plus la bijection réciproque :

$$f^{-1} : J \to I \;;\quad y \mapsto f^{-1}(y) = x \quad \text{tel que} \quad f(x) = y$$

est continue, strictement monotone, de même sens de variations que f. Dans un repère orhonormé, les courbes représentatives de f et f^{-1} sont symétriques par rapport à la droite d'équation $y = x$.

Application 1 : existence et étude d'une fonction comme réciproque d'une bijection.

Exemple de la fonction exponentielle.

On vérifie les hypothèses du théorème :

La fonction ln est continue et strictement croissante sur l'intervalle $]0\,;\,+\infty[$.

On donne les éléments nécessaires pour déterminer l'intervalle d'arrivée :

$$\lim_{x\to 0} \ln x = 0, \qquad \lim_{x\to +\infty} \ln x = +\infty.$$

On donne les conclusions, en citant le théorème utilisé :

Donc, d'après le théorème de la bijection monotone, La fonction ln est une bijection de $]0\,;\,+\infty[$ sur $]-\infty\,;\,+\infty[= \mathbf{R}$.

La bijection réciproque

$$\exp\ : \mathbf{R} \to]0\,;\,+\infty[\,, \quad y \mapsto \exp(y) = x \quad \text{tel que} \quad \ln x = y$$

est continue et strictement croissante sur $\mathbf{R}$.

$C_{\ln}$ et $C_{\exp}$ sont symétriques par rapport à D d'équation $y = x$.

À partir de ces éléments, on peut donner le tableau de variations de la fonction exponentielle.

Vous pouvez procéder au même travail de définition et d'étude de la fonction racine carrée.

Vous adapterez l'utilisation (fréquente) de ce théorème à la question posée.

Application 2 : existence de solutions à une équation. Une grande variété de situations est possible, donnons quelques exemples :

- Équation $f(x) = k$, k fixé dans $\mathbf{R}$. Soit $f(x) = e^x - x$. On a

x	$-\infty$		0		$+\infty$
$f(x)$	$+\infty$	↘	1	↗	$+\infty$

f réalise une bijection de $]-\infty\,;\,0]$ sur $[1\,;\,+\infty[$, donc, pour tout k dans $]1\,;\,+\infty[$, il existe un unique u_k négatif tel que $f(u_k) = k$. De même il existe un unique v_k positif tel que $f(v_k) = k$.

- Équation $f(x) = 0$. On peut utiliser, en l'adaptant, la proposition suivante :

> **Proposition.** Soit f une fonction continue et strictement monotone sur l'intervalle $[a\,;\,b]$, avec $f(a)f(b) < 0$. Alors l'équation $f(x) = 0$ admet une unique solution dans $]a\,;\,b[$.

Exemple. En étudiant f définie par $f(x) = x^3 - 3x^2 + 1$, on obtient

x	$-\infty$		0		2		$+\infty$
$f(x)$	$-\infty$	↗	1	↘	-3	↗	$+\infty$

On applique trois fois la proposition précédente :

Sur $]-\infty\,;\,0]$, f est continue et strictement croissante, $f(0) > 1$ et $\lim\limits_{-\infty} f = -\infty$, donc l'équation $f(x) = 0$ admet une unique solution. Rédaction analogue pour $]0\,;\,2[$ et $]2\,;\,+\infty[$.

On montre ainsi que l'équation $f(x) = 0$ admet exactement trois solutions a, b, c, avec $a < 0 < b < 2 < c$.

Méthode de dichotomie pour une valeur approchée de la solution.

On est dans les hypothèses de la proposition ci-dessus. Pour fixer les idées, on suppose f strictement croissante, et donc $f(a) < 0 < f(b)$.

On omet la partie déclarative du programme, y compris la déclaration de la fonction :

```
BEGIN
readln (a,b) ;
re eat
  c:=(a+b) 2 ;
  if f(c)>0 then b:=c  {la racine est entre a et c}
  else a:=c;{la racine est entre b et c}
until b-a <0.001;
writeln(a);{valeur a  rochée à 0,001  rès  ar défaut}
writeln(b);{à 0,001  rès  ar excès}
writeln(c);{à 5.10-4  rès}
END.
```

2.4 Image d'un segment par une fonction continue

Un **segment** de **R** est un intervalle fermé borné.

L'image d'un segment par une fonction continue est un segment.

Si $I = [a\,;\,b]$, on a donc, avec f continue sur I : $f(I) = [m\,;\,M]$.
m et M sont respectivement le **minimum** et le **maximum** de f sur I. On note

$$m = \min_{t\in[a\,;\,b]} f(t)\ ; \qquad M = \max_{t\in[a\,;\,b]} f(t)$$

La détermination de $f(I)$ se fait en utilisant le tableau de variations.

Exemple

x	-2		0		3		5
$f(x)$	5	↘	-3	↗	8	↘	3

En supposant f continue, l'image de l'intervalle $[-2\,;\,5]$ est l'intervalle $[-3\,;\,8]$ Sans cette supposition, on pourrait affirmer seulement $f([-2\,;\,5]) \subset [-3\,;\,8]$.

3. Calcul différentiel

3.1 Définitions – Opérations

Définitions

Soit f une fonction définie sur l'intervalle I, et $x_0 \in I$.

• f est dite **dérivable en** x_0 ssi il existe un nombre réel noté $f'(x_0)$ tel que :

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

• f est dite **dérivable sur** I ssi f est dérivable en tout point de I. La fonction

$$f' : I \to \mathbf{R}, \quad x_0 \mapsto f'(x_0)$$

est alors la **dérivée de** f **sur** I.

• Sous réserve d'existence, on définit la **dérivée seconde** de f : $f'' = \left(f'\right)'$. De façon générale, on définit, sous réserve d'existence, la **dérivée** n**-ème** de f par :

$$f^{(0)} = f\ ; \quad f^{(n)} = \left(f^{(n-1)}\right)' \text{ si } n \in \mathbf{N}^*$$

• Soit $n \in \mathbf{N}$. f est de dite **de classe C^n sur** I, ssi f est au moins n fois dérivable sur I, la dérivée n-ème $f^{(n)}$ étant continue sur I.

• f est dite **de classe C^∞ sur** I ssi f est indéfiniment dérivable sur I.

• Pour $n \in \mathbf{N} \cup \{\infty\}$, $\mathrm{C}^n(I)$ désigne l'ensemble des fonctions de classe C^n (ou : C^n) sur I. En particulier, $\mathrm{C}^0(I)$ est l'ensemble des fonctions continues sur I.

• On définit aussi la dérivée à droite en x_0 :

$$f'_{\mathrm{d}}(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}$$

sous réserve de limite finie. Définition analogue pour la dérivée à gauche.

Quelques propriétés :

• f dérivable en $x_0 \in]a, b[\Leftrightarrow f$ dérivable à gauche et à droite en x_0, et

$$f'_{\mathrm{d}}(x_0) = f'_{\mathrm{g}}(x_0).$$

• f dérivable sur $[a, b] \Leftrightarrow f$ dérivable sur $]a\,; b[$, dérivable à droite en a, dérivable à gauche en b.

• f est dérivable en x_0 ssi

$$f(x_0 + h) = f(x_0) + f'(x_0)\,h + h\varepsilon(h), \quad \text{avec } \lim_0 h = 0$$

(développement limité d'ordre 1 de f en x_0).

• Si f est dérivable en x_0, alors f est continue en x_0. Si f est C^1 sur I, alors f est dérivable sur I. Les réciproques sont fausses.

• $\mathrm{C}^0(I) \supset \mathrm{C}^1(I) \supset \mathrm{C}^2(I) \supset \ldots \supset \mathrm{C}^\infty(I)$.

Opérations

• Les fonctions polynômes, la fonction exponentielle, sont de classe C^∞ sur $\mathbf{R}$. Les fonctions ln, racine carrée, $x \mapsto x^r$ avec $r > 0$, sont de classe C^∞ sur $]0\,; +\infty[$.

Soit $n \in \mathbf{N} \cup \{\infty\}$.

• La somme, le produit, le quotient avec le dénominateur qui ne s'annule pas de deux fonctions C^n sur I sont des fonctions C^n sur I.

• Si u est C^n sur I et f est C^n sur $u(I)$, alors la composée $f \circ u$ est C^n sur I.

On montre qu'une fonction f est de classe C^n sur un intervalle, en disant que f est, suivant le cas, la somme, le produit... de fonctions de classe C^n. Attention au cas particulier $n = 1$, voir § 1.3.2.

3.2 Dérivée première

Calculs

a) Dérivées des fonctions usuelles

fonction	dérivée	commentaire
$x \mapsto b$	$x \mapsto 0$	sur $\mathbf{R}$
$x \mapsto x^n$ $x \mapsto x^r$	$x \mapsto nx^{n-1}$ $x \mapsto rx^{r-1}$	sur $\mathbf{R}$ si $n \in \mathbf{N}^*$, $\mathbf{R}^*$ si $n \in \mathbf{Z}^-$ sur $\mathbf{R}, \mathbf{R}^*, \mathbf{R}^{*+}$suivant la valeur de $r \in \mathbf{R}$
$x \mapsto \ln\|x\|$	$x \mapsto \frac{1}{x}$	sur $\mathbf{R}^*$
$x \mapsto e^x$	$x \mapsto e^x$	sur $\mathbf{R}$

b) Somme, produit, quotient : soit u et v dérivables sur I, $a \in \mathbf{R}$. Alors $u+v$, au, uv sont dérivables sur I, et

$$(u+v)' = u' + v' \quad ; \quad (au)' = au' \quad ; \quad (uv)' = u'v + uv'$$

Si de plus v ne s'annule pas sur I, alors $\frac{1}{v}$ et $\frac{u}{v}$ sont dérivables sur I, et

$$\left(\frac{1}{v}\right)' = -\frac{v'}{v^2} \quad ; \quad \left(\frac{u}{v}\right)' = \frac{u'v - v'u}{v^2}.$$

c) Composée : soit u dérivable sur I et f dérivable sur $u(I)$. Alors la composée $x \mapsto f(u(x))$ est dérivable sur I, de dérivée

$$x \mapsto u'(x) f'(u(x)).$$

d) Réciproque : si f est dérivable et f' ne s'annule pas sur l'intervalle I, alors f est une bijection dont la réciproque est dérivable sur l'intervalle $J = f(I)$, et on a, pour tout $y \in J$:

$$\left(f^{-1}\right)'(y) = \frac{1}{f'\left(f^{-1}(y)\right)}$$

Exemples de mise en œuvre :

- Utilisation de la dérivée de $x \mapsto x^r$:

$$f(x) = \frac{1}{x} \quad ; \quad f'(x) = -\frac{1}{x^2} \quad \text{pour} \quad x \neq 0$$

$$f(x) = \sqrt{x} \quad ; \quad f'(x) = \frac{1}{2\sqrt{x}} \quad \text{pour} \quad x > 0$$

(Formules d'emploi très fréquent.)

$$f(x) = \frac{1}{x^3} = x^{-3} \;;\quad f'(x) = -3x^{-4} = -\frac{3}{x^4},\ x \neq 0\,;$$

$$f(x) = x\sqrt{x} = xx^{\frac{1}{2}} = x^{\frac{3}{2}} \;;\quad f'(x) = \frac{3}{2}x^{\frac{1}{2}} = \frac{3}{2}\sqrt{x},\ x \geqslant 0.$$

• On retient les cas particuliers du c) :

$$(u^r)' = u' r u^{r-1} \;;\; \left(\sqrt{u}\right)' = \frac{u'}{2\sqrt{u}} \;;\; \left(\ln|u|\right)' = \frac{u'}{u} \;;\; (\mathrm{e}^u)' = u'\mathrm{e}^u$$

• Utilisation de la dérivée de u^r :

$$f(x) = \frac{1}{(1-x)^2} = (1-x)^{-2}\,;$$

$$f'(x) = (-1)(-2)(1-x)^{-3} = \frac{2}{(1-x)^3}$$

• On utilise le b) et le c) pour montrer qu'une fonction est dérivable (et calculer sa dérivée) sur un intervalle. En un point, on peut être amené à revenir à la définition :

$$f(x) = \frac{\ln x}{x - \ln x} \text{ si } x > 0 \;;\quad f(0) = -1.$$

Sur $]0\,;\,+\infty[$, f est dérivable car c'est le quotient de deux fonctions dérivables avec le dénominateur qui ne s'annule pas.

En 0, $\frac{f(x)-f(0)}{x-0} = \cdots = \frac{1}{x-\ln x} \underset{x\to 0}{\longrightarrow} 0$, donc f est dérivable en 0, et $f'(0) = 0$. (On pourrait aussi écrire $f'_{\mathrm{d}}(0) = 0$.)

Applications

Application au sens de variations d'une fonction

Soit f dérivable sur l'intervalle I.

Si $f' \geqslant 0$ (resp. $f' \leqslant 0$) sur I, alors f est croissante (resp. décroissante) sur I.

Si $f' > 0$ (resp. $f' < 0$) sur I sauf en un nombre fini de points, alors f est strictement croissante (resp. strictement décroissante) sur I.

Application aux tangentes à une courbe

Si f est dérivable en x_0, alors la courbe représentative C_f admet une tangente au point d'abscisse x_0. Cette tangente a pour coefficient directeur $f'(x_0)$ et a pour équation $y - f(x_0) = f'(x_0)(x - x_0)$.

Inégalités des accroissements finis

• Soit f une fonction dérivable sur $[a\,;b]$, avec $a \leqslant b$, telle que : $\forall x \in [a\,;b],\ m \leqslant f'(x) \leqslant M$. Alors :

$$m(b-a) \leqslant f(b) - f(a) \leqslant M(b-a)$$

• Soit f une fonction dérivable sur l'intervalle I telle que $\forall x \in I, |f'(x)| \leqslant k$. Alors, pour tout x_1, x_2 appartenant à I :

$$|f(x_1) - f(x_2)| \leqslant k|x_1 - x_2|$$

Dans la première inégalité des accroissements finis, on a $a < b$. Dans la deuxième inégalité, x_1 et x_2 sont dans un ordre quelconque. On parle aussi de « formule des accroissements finis ».

Exemple d'application : montrons que pour tout $n \in \mathbf{N}^*$,

$$\frac{1}{n+1} \leqslant \ln(n+1) - \ln(n) \leqslant \frac{1}{n}$$

• On repère quelle fonction f et quelles valeurs a, b on choisit, en comparant la formule générale et le résultat à obtenir. Ici, on voit que l'on doit prendre $f(x) = \ln x,\ a = n,\ b = n+1$.

• On encadre f' sur l'intervalle choisi. Ici :

$$f'(x) = \frac{1}{x}, \text{ donc } \frac{1}{n+1} \leqslant f'(x) \leqslant \frac{1}{n} \text{ pour tout } x \in [n\,;n+1]\,.$$

• On applique la formule, on vérifie que ça marche, puis on rédige.

Exemple de rédaction : soit $f(x) = \ln x, x > 0$. $\forall x \in [n\,;n+1]$, $\frac{1}{n+1} \leqslant f'(x) = \frac{1}{x} \leqslant \frac{1}{n}$,
donc, d'après la formule des accroissements finis :

$$\frac{1}{n+1}(n+1-n) \leqslant \ln(n+1) - \ln(n) \leqslant \frac{1}{n}(n+1-n)$$

$$\frac{1}{n+1} \leqslant \ln(n+1) - \ln(n) \leqslant \frac{1}{n}$$

Théorème du prolongement de la dérivée

Soit f continue sur $[a\,;b]$, de classe C^1 sur $]a\,;b]$ et telle que f' ait une limite finie ℓ en a.

Alors f est de classe C^1 sur $[a\,;b]$, et $f'(a) = \ell$.

Si f' a une limite infinie en a, alors f n'est pas dérivable en a. C_f admet une (demi-) tangente verticale au point d'abscisse a.

Pour établir que f' admet une limite finie en a, on pourra être amené à utiliser les développements limités :

$$f(x) = \frac{\ln(1+x)}{x} \quad \text{si} \quad x \in]0\,;\,+\infty[\quad ; \quad f(0) = 1.$$

• f est continue sur $]0\,;\,+\infty[$ (quotient de fonctions continues avec le dénominateur qui ne s'annule pas), et en 0 ($f(x)$ est équivalent en 0 à $\frac{x}{x} = 1$, donc $\lim\limits_{0} f = 1 = f(0)$), donc f est continue sur $[0\,;\,+\infty[$.

• f est de classe C^1 sur $]0\,;\,+\infty[$ comme quotient de fonctions de classe C^1 avec le dénominateur qui ne s'annule pas. Pour $x > 0$, on obtient

$$f'(x) = \frac{x - (1+x)\ln(1+x)}{x^2(1+x)}$$

• Pour passer à la limite $x \to 0$, on a une forme indéterminée, et on ne peut pas utiliser l'équivalent de $\ln(1+x)$ dans un somme. Mais on peut remplacer $\ln(1+x)$ par son DL en 0 (l'ordre 2 suffit) :

$$\begin{aligned} f'(x) &= \frac{x - (1+x)\left(x - \frac{x^2}{2} + x^2\varepsilon(x)\right)}{x^2\ln(1+x)} \\ &= \frac{x - \left(x - \frac{x^2}{2} + x^2\varepsilon(x) + x^2\right)}{x^2(1+x)} \\ &= \frac{-\frac{x^2}{2} + x^2\varepsilon(x)}{x^2(1+x)} = \frac{-\frac{1}{2} + \varepsilon(x)}{1+x} \end{aligned}$$

La limite ε en 0 est égale à 0, on obtient donc : $f'(x)$ tend vers $-\frac{1}{2}$ quand x tend vers 0.

• f est donc de classe C^1 sur $[0\,;\,+\infty[$, et $f'(0) = -\frac{1}{2}$.

3.3 Dérivée seconde

Définitions

Soit f une fonction dérivable sur l'intervalle I, $x_0 \in I$.

On dit que f est **convexe** sur I ssi la courbe C_f est au-dessus de ses tangentes en tout point de I.

On dit que f est **concave** ssi la courbe C_f est en-dessous de ses tangentes en tout point de I.

Le point $M_0\,(x_0, f\,(x_0))$ est un **point d'inflexion** de la courbe C_f ssi la tangente à C_f au point d'abscisse x_0 traverse C_f.
La tangente à C_f au point d'abscisse x ayant pour équation

$$y = f\,(x) + f'\,(x)\,(t - x)\,,$$

f est convexe sur I ssi, quels que soient x, t dans I :

$$f\,(t) \geqslant f\,(x) + f'\,(x)\,(t - x)\,.$$

Proposition

Pour f de classe C^2 sur l'intervalle I:

$$f \text{ convexe } \Leftrightarrow f'' > 0.$$
$$f \text{ concave } \Leftrightarrow f'' < 0.$$

Si f'' s'annule en changeant de signe en x_0, alors le point $M_0\,(x_0, f\,(x_0))$ est un point d'inflexion pour C_f.

Exemples d'applications :

• La fonction exponentielle est convexe sur $\mathbf{R}$. L'équation de la tangente à $C_{\exp}$ au point d'abscisse 0 est $y = x + 1$, on a donc :

$$\forall x \in \mathbf{R}, \quad e^x \geqslant x + 1$$

• La fonction ln est concave sur $]0\,;\,+\infty[$. L'équation de la tangente à $C_{\ln}$ au point d'abscisse 1 est $y = x - 1$, on a donc

$$\forall x > 0, \quad \ln x \leqslant x - 1$$

4. Fonctions usuelles

4.1 Fonctions exponentielle et logarithme népérien

Définitions. ln est la primitive de la fonction $x \mapsto \frac{1}{x}$ sur $]0\,;\,+\infty[$ qui s'annule en 1:

$$\forall x > 0,\ \ln'(x) = \frac{1}{x},\ \text{ et } \ln 1 = 0\ ;\ \text{ ou bien } \forall x > 0,\ \ln x = \int_1^x \frac{1}{t}\,dt$$

exp est la bijection réciproque de la fonction ln :

$$\exp\ : \mathbf{R} \to \mathbf{R}^{+*}, y \mapsto \exp(y) = x \quad \text{tel que} \quad \ln x = y$$

On note $\exp(y) = e^y$; $e \approx 2,718$.

Propriétés

Visibles sur le graphique (fig. 1) :

• ln est une bijection continue strictement croissante de $\mathbf{R}^{+*}$ sur $\mathbf{R}$.

$$\ln 1 = 0 \ ; \ \ln x < 0 \Leftrightarrow 0 < x < 1 \ ; \ \ln x > 0 \Leftrightarrow x > 1.$$

$$\lim_{x\to 0} \ln x = -\infty \ ; \ \lim_{x\to+\infty} \ln x = +\infty \ ; \ \lim_{x\to+\infty} \frac{\ln x}{x} = 0.$$

• exp est une bijection continue strictement croissante de $\mathbf{R}$ sur $\mathbf{R}^{+*}$.

$$e^0 = 1 \ ; \ e^x < 1 \Leftrightarrow x < 0 \ ; \ e^x > 1 \Leftrightarrow x > 0.$$

$$\lim_{x\to-\infty} e^x = 0 \ ; \ \lim_{x\to+\infty} e^x = +\infty \ ; \ \lim_{x\to+\infty} \frac{e^x}{x} = +\infty.$$

Autres propriétés analytiques :

• $\forall \alpha > 0, \ \lim\limits_{x\to 0} x^\alpha \ln x = 0 \ ; \ \lim\limits_{x\to+\infty} \dfrac{\ln x}{x^\alpha} = 0$

ln est C^∞ sur $\mathbf{R}^{+*}$; $\forall x > 0, \ \ln'(x) = \dfrac{1}{x}$; $\ln x \underset{1}{\sim} x - 1$.

• $\forall \alpha > 0, \ \lim\limits_{x\to+\infty} \dfrac{e^x}{x^\alpha} = +\infty.$

exp est C^∞ sur $\mathbf{R}$. $\forall x \in \mathbf{R}$, $\exp'(x) = \exp(x)$; $e^x - 1 \underset{0}{\sim} x$.

• $\forall x \in \mathbf{R}, \ \ln(e^x) = x$; $\forall x > 0, \ e^{\ln x} = x$.

Propriétés algébriques : ($\boldsymbol{a, b > 0}$; $\boldsymbol{x, y \in \mathbf{R}}$)

$\ln 1 = 0 \ ; \ \ln e = 1$	$e^0 = 1 \ ; \ e^1 = e$
$\ln(ab) = \ln a + \ln b$	$e^{x+y} = e^x e^y$
$\ln\left(\dfrac{1}{a}\right) = -\ln a$	$e^{-x} = \dfrac{1}{e^x}$
$\ln\left(\dfrac{a}{b}\right) = \ln a - \ln b$	$e^{x-y} = \dfrac{e^x}{e^y}$
$\ln(a^x) = x \ln a$	$(e^x)^y = e^{xy}$

4.2 Fonctions puissances

Définitions

• $x \in \mathbf{R}, n \in \mathbf{N}^*$: $x^n = x \times \cdots \times x$ (n facteurs)

• $x \in \mathbf{R}^*, n \in \mathbf{Z}^-$: $x^n = \frac{1}{x^{-n}}$ si $n < 0$; $x^0 = 1$

• $x \in \mathbf{R}^{+*}, p \in \mathbf{Z}, q \in \mathbf{N}^*$: $x^{\frac{p}{q}} = \sqrt[q]{x^p} = \left(\sqrt[q]{x}\right)^{\mathrm{p}}$
avec $\sqrt[q]{y}$ défini, pour $y \geqslant 0$, par $\sqrt[q]{y} \geqslant 0$ et $\left(\sqrt[q]{y}\right)^q = y$
En particulier, $x^{\frac{1}{2}} = \sqrt{x}$; $x^{\frac{1}{3}} = \sqrt[3]{x}$; $x^{-\frac{1}{2}} = \frac{1}{\sqrt{x}}$

• $x \in \mathbf{R}^{+*}, \quad r \in \mathbf{R} : \quad x^r = \mathrm{e}^{r \ln x}$

Règles de calcul

À chaque fois que toutes les écritures sont bien définies :

$$x^{r+r'} = x^r x^{r'} \;;\; x^{-r} = \frac{1}{x^r} \;;\; x^{r-r'} = \frac{x^r}{x^{r'}} \;;\; (x^r)^{r'} = x^{rr'}$$

$$\left(xx'\right)^r = x^r x'^r \;;\; \left(\frac{1}{x}\right)^r = \frac{1}{x^r} \;;\; \left(\frac{x}{x'}\right)^r = \frac{x^r}{x'^r}$$

Étude

• Avec $n \in \mathbf{N}^*$: $x \mapsto x^n$ est C^∞ sur $\mathbf{R}$, de dérivée $x \mapsto nx^{n-1}$, paire si n est pair, impaire si n est impair. Avec $p \in \mathbf{N}^*$:

x	$-\infty$		0		$+\infty$
x^{2p}	$+\infty$	$\searrow$	0	$\nearrow$	$+\infty$

x	$-\infty$		0		$+\infty$
x^{2p-1}	$-\infty$	$\nearrow$	0	$\nearrow$	$+\infty$

• Avec $r \in \mathbf{R} \setminus \mathbf{Z}$, $x \mapsto x^r$ est définie sur $\mathbf{R}^{+*}$, prolongeable par continuité en 0 si $r \geqslant 0$. Ce prolongement est de classe C^1 si $r \geqslant 1$.

Dans le cas général ($r \in \mathbf{R}$), les propriétés de croissance et de limite de la fonction $\mathbf{R}^+ \to \mathbf{R}^+$, $x \mapsto x^r$ se voient sur le graphique (**fig. 2**).

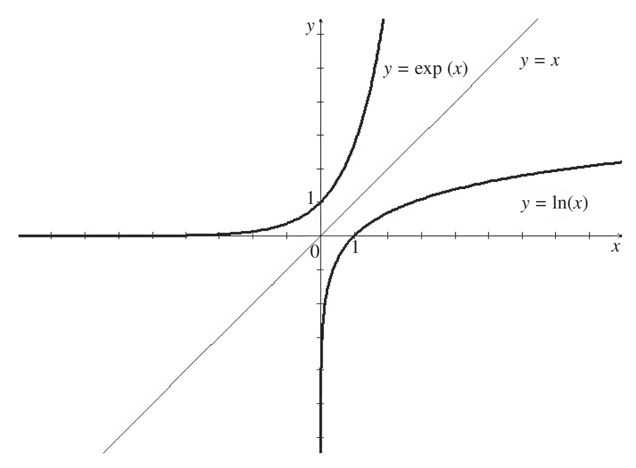


Fig. 1.1 Fonctions exponentielle et logarithme népérien

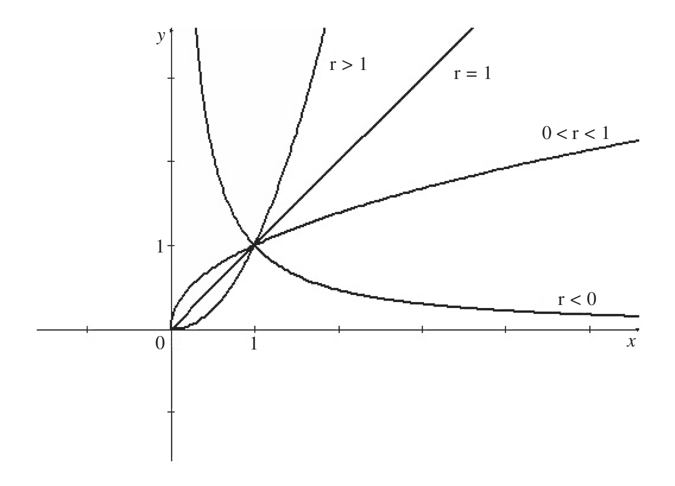


Fig. 1.2 Fonctions puissances $x \mapsto x^r$, $x > 0$

4.3 Fonction partie entière

La **partie entière** du réel x, notée Ent (x), ou $\lfloor x \rfloor$, est définie par :

$$\text{Ent}\,(x) = k \Leftrightarrow k \in \mathbf{Z} \quad \text{et} \quad k \leqslant x < k + 1$$

Ainsi, $\text{Ent}\,(3,5) = 3,\ \text{Ent}\,(-4,8) = -5,\ \text{Ent}\,(7) = 7.$

La fonction partie entière est continue à droite en tout point de $\mathbf{R}$; mais elle n'est pas continue à gauche en $k \in \mathbf{Z}$:

$$\lim_{x \longrightarrow k^-} \text{Ent}\,(x) = k - 1 \neq k = \text{Ent}\,(k)$$

La fonction Ent est donc continue sur $\mathbf{R} \setminus \mathbf{Z}$; l'ensemble de ses points de discontinuité est $\mathbf{Z}$.

La fonction **mantisse**, $m\,(x) = x - \text{Ent}\,(x)$, donne un exemple de fonction **périodique de période** 1 : $\forall x \in \mathbf{R}, m\,(x + 1) = m\,(x)$.

5. Fonctions de deux variables

5.1 Généralités

Éléments de topologie de $\mathbf{R}^2$. Soit $M_0 = (x_0, y_0)$ et $M = (x, y)$ deux points de $\mathbf{R}^2$. La **distance euclidienne** $\text{d}\,(M_0, M)$ est définie par

$$\text{d}\,(M_0, M) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

Un **ouvert** de $\mathbf{R}^2$ est une partie U de $\mathbf{R}^2$ telle que :

$$\forall M_0 \in U\,,\ \exists \alpha > 0\,,\ \text{d}\,(M_0, M) < \alpha \Rightarrow M \in U$$

Un **fermé** de $\mathbf{R}^2$ est une partie F de $\mathbf{R}^2$ telle que :

$$\left(\forall n \in \mathbf{N}, (x_n, y_n) \in F \ ; \ \lim_{n \to +\infty} x_n = x \ ; \ \lim_{n \to +\infty} y_n = y\right) \Rightarrow (x, y) \in F$$

Un ensemble **borné** de $\mathbf{R}^2$ est une partie B de $\mathbf{R}^2$ telle que :

$$\exists M \geqslant 0, \forall (x, y) \in B, \mathrm{d}\left((0,0), (x,y)\right) \leqslant M$$

Ces définitions n'ont pas à être mémorisées. Le fait que tel ensemble soit ouvert, fermé ou borné devrait vous être précisé.

Problème de l'optimisation. Soit D une partie de $\mathbf{R}^2$, et soit f une fonction de D dans $\mathbf{R}$. On cherche à déterminer les **extrema locaux**, ou **relatifs**, de f sur D, c'est-à-dire les valeurs $f(M_0)$ de f telles que :

$$\exists \alpha > 0, [M \in D \text{ et } \mathrm{d}(M_0, M) \leqslant \alpha \Rightarrow f(M_0) \geqslant f(M)]$$

(**maximum local**), ou

$$\exists \alpha > 0, [M \in D \text{ et } \mathrm{d}(M_0, M) \leqslant \alpha \Rightarrow f(M_0) \leqslant f(M)]$$

(**minimum local**).

Si l'inégalité a lieu pour tout $(x, y) \in D$, on parle d'extremum, de maximum, de minimum **global**, ou **absolu**.

Pour l'aide à la mémorisation, les résultats pour les fonctions d'une variable sont :

- Si f est continue sur l'intervalle fermé borné $[a\,;b]$, alors f admet un minimum et un maximum (globaux).
- Si f est de classe C^1 sur l'intervalle ouvert I, et si $f(x_0)$ est un extremum local, alors $f'(x_0) = 0$.
- Si f est de classe C^2 sur l'intervalle ouvert I, si $f'(x_0) = 0$ et si $f''(x_0) > 0$ (resp. $f''(x_0) < 0$), alors $f(x_0)$ est un minimum local (resp. un maximum local). (Attention au sens des inégalités.)

Dérivées partielles

Dérivées partielles d'ordre 1. Par définition, sous réserve de limite finie :

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}$$

On note aussi, respectivement : $f'_x(x_0, y_0)$, $f'_y(x_0, y_0)$.

Pratiquement, pour calculer $f'_x(x, y)$, on calcule la dérivée de f considérée comme une fonction de x, la variable y étant considérée comme une constante, et de même pour $f'_y(x, y)$:

$$f(x, y) = xy^2 + \frac{x}{y} + x^3 e^{-y} + y \quad \text{pour} \quad (x, y) \in \mathbf{R} \times \mathbf{R}^*$$

$$f'_x(x, y) = y^2 + \frac{1}{y} + 3x^2 e^{-y} \,;\, f'_y(x, y) = 2xy - \frac{x}{y^2} - x^3 e^{-y} + 1$$

Dérivées partielles d'ordre 2. On réitère le processus, ce qui conduit à 4 dérivées partielles d'ordre 2. En procédant méthodiquement :

$$\frac{\partial^2 f}{\partial x^2}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}(x, y)\right) \quad ; \quad \frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x, y)\right)$$

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x, y)\right) \quad ; \quad \frac{\partial^2 f}{\partial y^2}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}(x, y)\right)$$

On note aussi, respectivement, $f''_{x^2}(x, y)$, $f''_{yx}(x, y)$, $f''_{xy}(x, y)$, $f''_{y^2}(x, y)$.

Ainsi, pour l'exemple ci-dessus, on a :

$$f''_{x^2}(x, y) = 6xe^{-y} \quad ; \quad f''_{yx}(x, y) = 2y - \frac{1}{y^2} - 3x^2 e^{-y}$$

$$f''_{xy}(x, y) = 2y - \frac{1}{y^2} - 3x^2 e^{-y} \quad ; \quad f''_{y^2}(x, y) = 2x + \frac{2x}{y^3} + x^3 e^{-y}$$

5.2 Propriétés

Définitions. Soit D une partie non vide, ouverte ou fermée, de $\mathbf{R}^2$, et $f : D \to \mathbf{R}$.

On dit que :

- f est **continue en** $M_0 \in D$ ssi

$$\forall \varepsilon > 0, \exists \alpha > 0, M \in D \text{ et } d(M_0, M) < \alpha \Rightarrow d(f(M_0), f(M)) < \varepsilon$$

- f est **continue sur** D ssi f est continue en tout point de D.
- f est **de classe C¹ sur** D ssi les dérivées partielles d'ordre 1 de f existent et sont continues sur D.
- f est **de classe C² sur** D ssi les dérivées partielles d'ordre 2 de f existent et sont continues sur D.

L'ensemble des fonctions continues (resp. de classe C^1, de classe C^2), sur D est noté $C^0(D)$ (resp. $C^1(D)$, $C^2(D)$).

Opérations. Soit D une partie non vide, ouverte ou fermée, de $\mathbf{R}^2$, et $i \in \{0, 1, 2\}$.

a) Les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont de classe C^i sur $\mathbf{R}^2$.

b) Soit u, v appartenant à $C^i(D)$, $a \in \mathbf{R}$. Alors $u + v$, au et uv appartiennent à $C^i(D)$. $\frac{u}{v}$ appartient à $C^i(D)$ si de plus v ne s'annule pas sur D.

c) Soit u de classe C^i et ϕ de classe C^i sur $u(D)$. Alors la composée $\varphi \circ u$ est de classe C^i sur D.

Théorème de Schwarz.

$$\text{Si } f \in C^2(D), \quad \text{alors} \quad \frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y)$$

Notations de Monge. Pour $f \in C^2(D)$, on note, de façon abrégée :

$$p = f'_x \;; \quad q = f'_y \;; \quad r = f''_{x^2} \;; \quad s = f''_{yx} = f''_{xy} \;; \quad t = f''_{y^2}$$

5.3 Résultats d'optimisation

• Une condition suffisante d'extremum global :

Soit f une fonction **continue** sur une partie **fermée bornée** de $\mathbf{R}^2$. Alors f admet un minimum et un maximum globaux sur D.

• Une condition nécessaire d'extremum :

Soit f **de classe C^1** sur l'**ouvert** U de $\mathbf{R}^2$, et $M_0 = (x_0, y_0) \in U$. Si f présente un extremum en M_0, alors

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

On dit que (x_0, y_0) est un **point critique** de f.

• Une condition suffisante d'extremum :

Soit f **de classe C^2** sur l'**ouvert** U de $\mathbf{R}^2$, et $M_0 = (x_0, y_0) \in U$. Si, au point M_0, on a

• $p = q = 0$; $rt - s^2 > 0$; $r < 0$ (ou $t < 0$), alors $f(M_0)$ est maximum local de f.

• $p = q = 0$; $rt - s^2 > 0$; $r > 0$ (ou $t > 0$), alors $f(M_0)$ est minimum local de f.

Si $p = q = 0$; $rt - s^2 < 0$, alors $f(M_0)$ n'est pas un extremum local de f.

Dans le cas où $p = q = 0$; $rt - s^2 = 0$, on ne peut rien conclure.

Ces résultats s'établissent à partir du développement limité d'ordre 1 de f en (x_0, y_0), valable pour $f \in C^1(U)$, $(x_0, y_0) \in U$:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + ph + qk + \sqrt{h^2 + k^2}\ \varepsilon(h, k),$$

et à partir du développement limité d'ordre 2 de f en (x_0, y_0), valable pour $\in C^2(U)$, $(x_0, y_0) \in U$:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + ph + qk + \frac{1}{2}\left(rh^2 + 2shk + tk^2\right) + \left(h^2 + k^2\right)\varepsilon(h, k),$$

avec p, q, r, s, t pris au point (x_0, y_0), et avec ε continue et nulle au point $(0, 0)$.

Suites et séries numériques

2

1. Généralités

1.1 Définitions

Une **suite numérique** est une application de $\mathbf{N}$, ou d'une partie de $\mathbf{N}$, dans $\mathbf{R}$.

Si u est une suite numérique, au lieu de $u(n)$, on préfère écrire u_n (lire « u indice n »). u_n est appelé le **terme de rang** n de la suite u.

La suite u elle-même est notée $(u_n)_{n\in\mathbf{N}}$ (si elle est définie sur $\mathbf{N}$), ou simplement (u_n) si il n'y a pas d'ambiguïté sur l'ensemble de départ.

Une suite est un cas particulier de fonction numérique ; on retrouve le même vocabulaire, et en adaptant les définitions on a :

La suite $u = (u_n)_{n\in\mathbf{N}}$ est dite

- **croissante**, resp. **décroissante** ssi

$$\forall n \in \mathbf{N},\quad u_n \leqslant u_{n+1},\qquad \text{resp. } \forall n \in \mathbf{N},\quad u_n \geqslant u_{n+1}\ ;$$

- **monotone** ssi u est croissante ou décroissante ;
- **majorée** par M, resp. **minorée** par m ssi

$$\forall n \in \mathbf{N},\quad u_n \leqslant M,\qquad \text{resp. } \forall n \in \mathbf{N},\quad u_n \geqslant m\ ;$$

M est alors un **majorant**, resp. m est un **minorant** de u ;

- **bornée** ssi u est majorée et minorée.

Pour les suites, le seul problème de limite qui se pose est la limite de u_n quand n tend vers $+\infty$:

La suite (u_n) est dite **convergente** ssi il existe $\ell \in \mathbf{R}$ tel que

$$\forall \varepsilon > 0, \exists n_0 \in \mathbf{N},\quad n \geqslant n_0 \Rightarrow |u_n - \ell| \leqslant \varepsilon$$

Le nombre réel est alors appelé la **limite** de la suite (u_n), et on dit que la suite (u_n) **converge** vers ℓ. On dira souvent : « (u_n) tend vers ℓ ».

(u_n) converge vers ℓ ssi, pour tout $\varepsilon > 0$, il n'y a qu'un nombre fini de termes de la suite en dehors de l'intervalle $]\ell - \varepsilon, \ell + \varepsilon[$.

Si (u_n) converge vers ℓ et si $a < \ell < b$, alors, pour tous les termes de la suite à partir d'un certain rang : $a < u_n < b$.

1.2 Théorèmes de convergence

Suites monotones

Théorème 1

Une suite croissante et majorée est convergente.

Une suite décroissante et minorée est convergente.

Théorème admis.

En utilisant le passage à la limite dans les inégalités (voir § 1.1.2 ou ci-dessous, § 2.1.3), on peut préciser :

Si une suite est croissante et majorée par M, alors elle est convergente, et sa limite ℓ vérifie $\ell \leqslant M$. Si une suite croissante n'est pas convergente, alors $\lim\limits_{n \to +\infty} u_n = +\infty$.

Si une suite est décroissante et minorée par m, alors elle est convergente, et sa limite ℓ vérifie $\ell \geqslant m$. Si une suite décroissante n'est pas convergente, alors $\lim\limits_{n \to +\infty} u_n = -\infty$.

Soit x un nombre réel fixé dans $]0, 1[$, et soit (u_n) la suite définie par

$$\forall n \in \mathbf{N}, \quad u_n = \prod_{k=0}^{n} \left(1 + x^k\right)$$

La suite (u_n) est **croissante** car, pour tout $n \in \mathbf{N}$, $u_n \geqslant 0$ et $1+x^{n+1} \geqslant 1$, donc $u_{n+1} = u_n \left(1 + x^{n+1}\right) \geqslant u_n$

La suite (u_n) est **majorée**. Pour tout $n \in \mathbf{N}$, $u_n > 0$, et on a

$$\ln(u_n) = \sum_{k=0}^{n} \ln\left(1 + x^k\right) \leqslant \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} \leqslant \frac{1}{1 - x}$$

(Règles de calcul sur les logarithmes, puis utilisation de l'inégalité $\ln(1 + y) \leqslant y$, valable pour tout $y > -1$ et que l'on établit en étudiant la fonction $y \mapsto \ln(1 + y) - y$, puis identité géométrique ; la dernière majoration, évidente car $0 < x < 1$, est indispensable car le majorant

de la suite ne doit pas dépendre de n.) On a donc

$$\forall n \in \mathbf{N}, \quad u_n \leqslant e^{\frac{1}{1-x}}$$

La suite (u_n) est donc **croissante** et **majorée** par $M = e^{\frac{1}{1-x}}$. Elle est donc **convergente**, et sa limite ℓ vérifie $\ell \leqslant M$.

• Ce théorème s'applique si la suite est monotone à partir d'un certain rang.

• Ce théorème donne une condition suffisante, mais non nécessaire, pour qu'une suite soit convergente. En d'autres termes, il existe des suites convergentes qui ne sont ni croissantes ni décroissantes, par exemple la suite $(u_n)_{n \in \mathbf{N}^*}$ définie par $\forall n \in \mathbf{N}^*$, $u_n = \frac{(-1)^n}{n}$, qui converge vers 0.

• Pour montrer qu'une suite est (par exemple) croissante, les techniques courantes sont :

– Appliquer la définition.
– Montrer que $u_{n+1} - u_n \geqslant 0$; à utiliser si u_n se présente sous forme de somme, voir § 2.4.3 « séries ».
– Si les termes de la suite (u_n) sont positifs, montrer que $\frac{u_{n+1}}{u_n} \geqslant 1$; à utiliser si u_n se présente sous forme de produit, comme dans l'exemple précédent : avec $x > 0$, on a

$$\forall n \in \mathbf{N}, \quad u_n = \prod_{k=0}^{n} \left(1 + x^k\right) > 0 \,; \qquad \frac{u_{n+1}}{u_n} = 1 + x^{n+1} > 1$$

donc la suite (u_n) est (strictement) croissante.
– Pour une suite du type $u_{n+1} = f(u_n)$ (§ 2.2.3) ou pour une suite définie implicitement (§ 2.2.5), voir les techniques spécifiques dans les paragraphes suivants.

• Pour montrer qu'une suite est majorée (par exemple), on fera grand usage des manipulations des inégalités, voir le § 8 de l'introduction.

Suites adjacentes

Définition. Les deux suites $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ sont dites **adjacentes** ssi une des suites est croissante, l'autre décroissante, et

$$\lim_{n \to +\infty} (u_n - v_n) = 0$$

Théorème. Si deux suites sont adjacentes, alors elles sont convergentes et elles ont même limite.

Soient (u_n) et (v_n) deux suites adjacentes, de limite commune ℓ, la suite (u_n) étant croissante et la suite (v_n) décroissante. On a alors

$$u_0 \leqslant \cdots \leqslant u_n \leqslant \ell \leqslant v_n \leqslant \cdots \leqslant v_0$$

u_n est donc une valeur approchée par défaut, et v_n une valeur approchée par excès, de ℓ à moins de $v_n - u_n$ près.

Soit (u_n) et (v_n) les suites définies par $(n \in \mathbf{N}^*)$:

$$u_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln(n) \ ; \qquad v_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln(n+1)$$

Les suites (u_n) et (v_n) sont adjacentes. En effet, la suite (u_n) est décroissante, car pour tout $n \geqslant 1$,

$$u_{n+1} - u_n = \frac{1}{n+1} - \ln(n+1) + \ln(n) \leqslant 0$$

d'après la formule des accroissements finis, appliquée à la fonction ln sur l'intervalle $[n, n+1]$. On démontre de même que la suite (v_n) est croissante. La différence $(u_n - v_n)$ converge vers 0, en effet

$$u_n - v_n = \ln(n+1) - \ln(n) = \ln\left(\frac{n+1}{n}\right) = \ln\left(1 + \frac{1}{n}\right)$$

La limite commune aux suites (u_n) et (v_n) est notée γ (constante d'Euler). Pour tout n, on a $u_n \leqslant \gamma \leqslant v_n$; u_n est donc une valeur approchée par défaut de γ à moins de $v_n - u_n = \ln\left(\frac{n}{n+1}\right)$ près.

D'autre part, en posant $\varepsilon_n = u_n - \gamma$, on obtient :

$$1 + \frac{1}{2} + \cdots + \frac{1}{n} = \ln(n) + \gamma + \varepsilon_n$$

avec (ε_n) qui converge vers 0. Cela montre en particulier que la suite de terme général $1 + \frac{1}{2} + \cdots + \frac{1}{n}$ tend vers $+\infty$ (résultat à retenir, voir § 2.4.3).

1.3 Opérations sur les limites

Une suite est cas particulier de fonction numérique : l'ensemble de définition de la suite est $\mathbf{N}$, ou une partie de $\mathbf{N}$. On peut donc utiliser toutes les connaissances et techniques du chapitre 1, grâce à la proposition suivante :

Soit f une application définie sur $[0, +\infty[$.

$$\text{Si } \lim_{x \to +\infty} f(x) = \ell, \quad \text{alors } \lim_{n \to +\infty} f(n) = \ell$$

On reprend les points essentiels :

• Limites classiques

Avec $r > 0$: $\lim\limits_{n\to+\infty} n^r = +\infty$; $\lim\limits_{n\to+\infty} \frac{1}{n^r} = 0$;

$\lim\limits_{n\to+\infty} \ln(n) = +\infty$

Avec $x > 1$: $\lim\limits_{n\to+\infty} x^n = +\infty$

Avec $-1 < x < 1$: $\lim\limits_{n\to+\infty} x^n = 0$

• Négligeabilités classiques

Avec $\alpha > 0,\ x > 1$: $\lim\limits_{n\to+\infty} \frac{\ln(n)}{n^\alpha} = 0$; $\lim\limits_{n\to+\infty} \frac{x^n}{n^\alpha} = +\infty$

Avec $\alpha > 0,\ -1 < x < 1$: $\lim\limits_{n\to+\infty} n^\alpha x^n = 0$

• Limite d'une somme, d'un produit, d'un quotient

Soit (u_n) et (v_n) telles que $\lim\limits_{n\to+\infty} u_n = \ell,\ \lim\limits_{n\to+\infty} v_n = \ell'$. Alors

$$\lim_{n\to+\infty} (u_n + v_n) = \ell + \ell' \; ; \; \lim_{n\to+\infty} (u_n v_n) = \ell\ell'$$

$$\lim_{n\to+\infty} \left(\frac{u_n}{v_n}\right) = \frac{\ell}{\ell'} \text{ si de plus } \ell' \neq 0$$

• Limite de $f(u_n)$

$$\text{Si } \lim_{n\to+\infty} u_n = \ell \text{ et si } \lim_{\ell} f = L, \text{ alors } \lim_{n\to+\infty} f(u_n) = L$$

En particulier :

$$\text{Si } \lim_{n\to+\infty} u_n = \ell \in \mathbf{R} \text{ et si } f \text{ est continue en } \ell, \text{ alors}$$

$$\lim_{n\to+\infty} f(u_n) = f(\ell)$$

• Passage à la limite dans les inégalités

$\forall n \geqslant n_0, u_n \leqslant v_n$; $\lim\limits_{n\to+\infty} u_n = \ell$; $\lim\limits_{n\to+\infty} v_n = \ell' \Rightarrow \ell \leqslant \ell'$

$\forall n \geqslant n_{0,}, u_n \leqslant v_n \leqslant w_n$; $\lim\limits_{n\to+\infty} u_n = \lim\limits_{n\to+\infty} w_n = \ell \Rightarrow \lim\limits_{n\to+\infty} v_n = \ell$

$\forall n \geqslant n_0\ |u_n - \ell| \leqslant v_n$; $\lim\limits_{n\to+\infty} v_n = 0 \Rightarrow \lim\limits_{n\to+\infty} u_n = \ell$

• **Équivalents.** En adaptant la définition générale, on dit que (u_n) et (v_n) sont équivalentes ssi $\lim\limits_{n\to+\infty} \frac{u_n}{v_n} = 1$. On a les mêmes propriétés, et la même utilisation pour la recherche de limite.

> Quand u_n tend vers $0, +\infty$ ou $-\infty$, l'obtention d'un équivalent pour u_n permet d'apprécier qualitativement le comportement de u_n :
> On a obtenu dans l'exemple donné pour les suites adjacentes :
> $$1 + \frac{1}{2} + \cdots + \frac{1}{n} = \ln(n) + \gamma + \varepsilon_n$$
> avec γ une constante et (ε_n) qui converge vers 0. En divisant cette égalité par $\ln(n)$, on obtient
> $$1 + \frac{1}{2} + \cdots + \frac{1}{n} \underset{n\to+\infty}{\sim} \ln(n)$$

• **Négligeabilité.** En adaptant ici aussi la définition générale, on dit que u_n est négligeable devant v_n, et on note $u_n = o(v_n)$, ssi
$$\lim_{n\to+\infty} \frac{u_n}{v_n} = 0$$

2. Suites numériques calculables

2.1 Suites arithmétiques

Définition. Soit $r \in \mathbf{R}$. La suite $(u_n)_{n\in\mathbf{N}}$ est dite **arithmétique de raison** r ssi
$$\forall n \in \mathbf{N}, \quad u_{n+1} = u_n + r.$$

On démontre alors par récurrence la

> **Proposition 1.** La suite $(u_n)_{n\in\mathbf{N}}$ est arithmétique de raison r ssi
> $$\forall n \in \mathbf{N}, \quad u_n = u_0 + nr$$

On parle également de suite arithmétique définie sur $\mathbf{N}^*$:
$\forall n \in \mathbf{N}^*, \ u_{n+1} = u_n + r$. On a alors : $\forall n \in \mathbf{N}^*, \ u_n = u_1 + (n-1)r$.
De façon générale, pour tout n, p tels que $u_n$et u_p soient définis :
$$u_n = u_p + (n-p)r$$

> **Proposition 2.** Pour une suite arithmétique (u_n), on a
> $$\sum_{k=1}^{n} u_k = n \times \frac{u_1 + u_n}{2}$$

En effet,

$$\begin{aligned} S_n &= u_1 + u_2 + \cdots + u_n \\ &= u_n + u_{n-1} + \cdots + u_1, \end{aligned}$$

et donc $2S_n = (u_1 + u_n) + (u_1 + r + u_n - r) + \cdots + (u_n + u_1)$

$= n\,(u_1 + u_n)$

Avec $u_k = k$, on obtient le cas particulier important, à retenir :

$$\sum_{k=1}^{n} k = \frac{n\,(n+1)}{2}$$

De façon générale, la somme de termes successifs d'une suite arithmétique est égale au nombre de termes multiplié par la moyenne arithmétique des termes extrêmes.

Le sens de variations et la limite d'une suite arithmétique de raison r ne pose aucun problème : Si $r > 0$, la suite est strictement croissante, tend vers $+\infty$; Si $r < 0$, la suite est strictement décroissante, tend vers $-\infty$. La suite est constante si $r = 0$;

2.2 Suites géométriques

Définition. Soit r un nombre réel non nul. La suite $(u_n)_{n\in\mathbf{N}}$ est dite **géométrique de raison** r ssi :

$$\forall n \in \mathbf{N}, \quad u_{n+1} = ru_n$$

Théorème 1. La suite $(u_n)_{n\in\mathbf{N}}$ est géométrique de raison r ssi

$$\forall n \in \mathbf{N}, \quad u_n = u_0 r^n$$

Dans le cas particulier où $u_n = x^n$, avec $x \neq 1$, on a le très important

Théorème 2. Pour tout $n \in \mathbf{N}$, et pour tout $x \neq 1$:

$$\sum_{k=0}^{n} x^k = 1 + x + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}$$

Pour le sens de variations et les limites, on a :

Théorème 3.

- Si $x > 1$, $\lim_{n\to+\infty} x^n = +\infty$.

Si $-1 < x < 1$, $\lim_{n\to+\infty} x^n = 0$.

Si $x \leqslant -1$, la suite (x^n) n'a pas de limite.

- Si $x > 1$, la suite (x^n) est strictement croissante.

Si $0 < x < 1$, la suite (x^n) est strictement décroissante.

Si $x < 0$, la suite (x^n) n'est ni croissante ni décroissante.

Le théorème 1 est tout à fait fondamental et d'usage constant, voir les paragraphes suivants par exemple. De façon typique, sa démonstration se fait par récurrence : la propriété à établir est vraie pour $n = 0$, car $u_0 r^0 = u_0$, et si il existe $n \in \mathbf{N}$ tel que $u_n = u_0 r^n$, alors $u_{n+1} = ru_n = ru_0 r^n = u_0 r^{n+1}$, d'où la conclusion.

On parle également de suite géométrique définie sur $\mathbf{N}^*$:

$\forall n \in \mathbf{N}^*,\ u_{n+1} = ru_n$. On a alors : $\forall n \in \mathbf{N}^*,\ u_n = u_1 r^{n-1}$.

De façon générale, pour tout n, p tels que u_n et u_p soient définis :

$$u_n = u_p r^{n-p}$$

La formule du théorème 2 est connue sous le nom d'**identité géométrique**. Elle est valable pour tout $x \neq 1$. Avec $x = 1$, on obtient bien sûr une somme égale à $n + 1$, car les $n + 1$ termes de la somme sont tous égaux à 1.

En mettant en facteur le premier terme, vous pouvez calculer grâce au théorème 2 la somme de termes successifs d'une suite géométrique, et vous pouvez retenir la formule (valable si la raison est $\neq 1$) :

$$\text{Somme de termes successifs d'une suite géométrique} = \text{premier terme} \times \frac{1 - \text{raison}^{\text{nb de termes}}}{1 - \text{raison}}$$

Voici quelques usages typiques de l'identité géométrique. n est un entier naturel, non nul dans la première et la dernière formule, x un nombre réel différent de 1 et -1 :

- $\sum_{k=1}^{n} 2^k = 2^n - 1$; $\quad \sum_{k=0}^{n} \left(\frac{1}{2}\right)^k = 2\left(1 - \left(\frac{1}{2}\right)^{n+1}\right)$

$$\bullet \sum_{k=0}^{n} \mathrm{e}^{-k} = \sum_{k=0}^{n} \left(\frac{1}{\mathrm{e}}\right)^k = \frac{\left(1 - \left(\frac{1}{\mathrm{e}}\right)^{\mathrm{n}+1}\right)}{1 - \frac{1}{\mathrm{e}}} = \frac{\mathrm{e}}{\mathrm{e}-1}\left(1 - \left(\frac{1}{\mathrm{e}}\right)^{n+1}\right)$$

$$\bullet \sum_{k=0}^{n} x^{2k} = \sum_{k=0}^{n} \left(x^2\right)^k = \frac{1-\left(x^2\right)^{n+1}}{1-x^2} = \frac{1-x^{2n+2}}{1-x^2}$$

$$\bullet \sum_{k=1}^{n} x^{2k+1} = x \sum_{k=1}^{n} \left(x^2\right)^k = xx^2 \frac{1-\left(x^2\right)^n}{1-x^2} = x^3 \frac{1-x^{2n}}{1-x^2}$$

Veillez à ne pas confondre par exemple les sommes $\sum_{k=1}^{n} 4^k$ (susceptible d'être calculée avec l'identité géométrique), et $\sum_{k=1}^{n} k^4$ (qui ne se calcule pas de manière simple).

On mémorise les résultats du théorème 3 en prenant des cas particuliers :

$$x = 2, \quad x = \frac{1}{2}, \quad x = -\frac{1}{2}, \quad x = -2.$$

Dans le théorème 3, vous retiendrez particulièrement les deux premiers résultats. Ils se démontrent en utilisant les logarithmes ; par exemple :

Si $0 < |x| < 1$, $|x^n| = |x|^n = \mathrm{e}^{n\ln|x|} \underset{x\to+\infty}{\longrightarrow} 0$ car $\ln|x| < 0$.

2.3 Suites arithmético-géométriques

Théorème. Soit $a \neq 1$, $b \in \mathbf{R}$, $(u_n)_{n\in\mathbf{N}}$ telles que

$$\forall n \in \mathbf{N}, \quad u_{n+1} = au_n + b.$$

Alors, pour tout $n \in \mathbf{N}$:

$$u_n - \ell = a^n (u_0 - \ell),$$

avec ℓ tel que

$$\ell = a\ell + b.$$

En effet, ℓ existe car on a supposé $a \neq 1$. En soustrayant membre à membre les deux égalités $u_{n+1} = au_n + b$, $\ell = a\ell + b$, il vient $u_{n+1} - \ell = a(u_n - \ell)$, ce qui prouve que la suite $(u_n - \ell)$ est géométrique de raison a, d'où le résultat, en utilisant le théorème 1 du paragraphe précédent.

Si la suite arithmético-géométrique est définie sur $\mathbf{N}^*$, le théorème s'adapte facilement, et on trouve $u_n - \ell = a^{n-1}(u_1 - \ell)$.

2.4 Suites linéaires récurrentes à deux termes

Théorème. Soit $a, b \in \mathrm{R}$, $(u_n)_{n\in\mathrm{N}}$ telles que

$$\forall n \in \mathbf{N}, \quad u_{n+2} = au_{n+1} + bu_n$$

On considère l'équation caractéristique (1) $\quad r^2 = ar + b$.

• Si l'équation (1) a deux racines r_1, r_2, alors il existe $\alpha, \beta \in \mathrm{N}$ tels que :

$$\forall n \in \mathbf{N}, \quad u_n = \alpha r_1^n + \beta {r_2}^n$$

• Si l'équation (1) a une racine double r_1, alors il existe $\alpha, \beta \in \mathrm{N}$ tels que :

$$\forall n \in \mathbf{N}, \quad u_n = \alpha {r_1}^n + \beta n {r_1}^n$$

Pour déterminer α et β, on a besoin d'informations supplémentaires sur la suite (u_n), par exemple la donnée de u_0 et u_1.

La démonstration de ce théorème est un bon exemple de l'utilisation des outils de l'algèbre linéaire, voir § 5.5.1.

Programmation du cas particulier de la suite de Fibonacci :

$$u_0 = 1\,;\; u_1 = 1\,;\; \forall n \in \mathbf{N}, \quad u_{n+2} = u_{n+1} + u_n$$

On doit écrire un algorithme itératif. Les appels multiples dans un algorithme récursif donnent des résultats non souhaités.

Voir § 10.1.8.

Le programme suivant affiche les valeurs de $u_2, \ldots, u_{100}$:

```
program fibonacci ;
var i:integer ; a,b,c:real ;
BEGIN
a:=1 ;b:=1 ;
for i:=2 to 100 do begin
c:=a+b ;
 writeln(i,c) ; {écriture de u_i avec son rang i}
 a:=b ; b:=c ; {on prépare l'éventuelle itération suivante}
 end ;
END.
```

3. Suites $u_{n+1} = f(u_n)$

3.1 Généralités

Définition. Dans tout le paragraphe, (u_n) désigne une suite telle que :

$$\forall n \in \mathbf{N}, \quad u_{n+1} = f(u_n)$$

où f est une fonction de $\mathbf{R}$ dans $\mathbf{R}$.

Attention, les résultats donnés ne sont valables que pour les suites de ce type ! On adaptera sans peine ces résultats au cas où la suite est définie sur $\mathbf{N}^*$.

En général, l'énoncé admet implicitement que la suite (u_n) est bien définie (c'est évident si f est définie sur $\mathbf{R}$). Si la question est posée, on répondra au moyen d'un raisonnement par récurrence.

Soit (u_n) définie par :

$$u_0 = 1\,; \;\; \forall n \in \mathbf{N}, \quad u_{n+1} = u_n + \frac{1}{u_n}$$

On montre par récurrence : $\forall n \in \mathbf{N}, u_n$ existe et $u_n > 0$; $u_0 = 1$ existe et $u_0 > 0$, et si u_n existe et $u_n > 0$, alors $u_{n+1} = u_n + \frac{1}{u_n}$ existe et $u_{n+1} > 0$. Ce qui montre que la suite (u_n) est bien définie.

Sens de variation

Théorème. Soit I un intervalle de $\mathbf{R}$.
Si $\forall n \in \mathbf{N}, u_n \in I$, et f est **croissante** sur I, alors $(u_n)_{n\in\mathbf{N}}$ est **monotone**.

On démontre ce théorème par récurrence, en distinguant les cas $u_0 \leqslant u_1$, $u_0 \geqslant u_1$. Si $u_0 \leqslant u_1$, on démontre : $\forall n \in \mathbf{N}, u_n \leqslant u_{n+1}$. La propriété est vraie pour $n = 0$ par hypothèse, et si elle vraie pour n fixé dans $\mathbf{N}$, alors $u_n \leqslant u_{n+1}$, donc $f(u_n) \leqslant f(u_{n+1})$ car f est croissante, donc $u_{n+1} \leqslant u_{n+2}$, et la propriété est vraie pour $n + 1$. On en conclut qu'elle est vaie pour tout n, ce qui montre que la suite (u_n) est croissante. On procède de même si $u_0 \geqslant u_1$.

Ce théorème est à connaître, mais il existe d'autres moyens pour prouver la monotonie de (u_n). Par exemple, si on sait que :

$$\forall x \in I, \;\; f(x) \geqslant x\,; \qquad \forall n \in \mathbf{N}, \quad u_n \in I$$

on obtient, sans utiliser le théorème, que (u_n) est croissante, puisqu'on a alors :

$$\forall n \in \mathbf{N}, \quad u_{n+1} = f(u_n) \geqslant u_n$$

Remarque. Si f est décroissante sur I, avec $u_n \in I$ pour tout n, la suite (u_n) n'est ni croissante ni décroissante, mais les suites (u_{2n}) et(u_{2n+1}) sont monotones, car $f \circ f$ est croissante, et

$$u_{2(n+1)} = f \circ f\,(u_{2n})\,, \quad u_{2(n+1)+1} = f \circ f\,(u_{2n+1})$$

Si f n'est ni croissante ni décroissante, le comportement de (u_n) peut s'avérer très complexe. On peut en faire une approche numérique.

Notion de point fixe

Définition. On dit que $a \in \mathbf{R}$ est un **point fixe** de f ssi $f(a) = a$.

> **Théorème.** Si (u_n) converge vers ℓ et f est continue en ℓ, alors ℓ est un point fixe de f, c'est-à-dire que $\ell = f(\ell)$.

On obtient ce résultat par passage à la limite dans $u_{n+1} = f(u_n)$.

En effet, si $\lim\limits_{n\to+\infty} u_n = \ell$, alors $\lim\limits_{n\to+\infty} u_{n+1} = \ell$, et $\lim\limits_{n\to+\infty} f(u_n) = f(\ell)$ car f est continue en ℓ

Remarque. Ce théorème donne une condition nécessaire, mais non suffisante, pour que (u_n) converge vers ℓ. Le problème est que ℓ est inconnue, donc on ne sait pas *a priori* si f est continue en ℓ ! On a néanmoins dans cette direction :

• On suppose que f continue sur $\mathbf{R}$. Alors, si (u_n) converge vers ℓ, ℓ est un point fixe de f.

• On suppose que f continue sur l'intervalle $I = [a, b]$, et que tous les termes de la suite (u_n) appartiennent à I. Alors, si (u_n) converge vers ℓ, ℓ est un point fixe de f.

En effet, $a \leqslant u_n \leqslant b$ pour tout n, donc si (u_n) converge vers ℓ, alors $a \leqslant \ell \leqslant b$ (passage à la limite dans les inégalités). ℓ appartient donc à I, f est continue en ℓ, et ℓ est un point fixe de f. On peut faire le même raisonnement avec tout intervalle I **fermé**.

3.2 Exemples

Une grande variété de situations est possible. Vous répondrez aux questions posées, en cherchant à comprendre leur enchaînement logique. Donnons quelques indications générales :

On montre que la suite (u_n) est majorée, ou minorée, ou bornée, en général par récurrence, en utilisant les points fixes de f.

L'énoncé propose souvent l'étude du signe de $f(x) - x$, qui fournit les points fixes de f (avec $f(x) - x = 0$), ou les intervalles I tels

que, si tous les u_n appartiennent à I, alors (u_n) est décroissante (avec $f(x) - x \leqslant 0$), ou croissante (avec $f(x) - x \geqslant 0$).

• $u_0 = a > 0\,;\ \forall n \in \mathbf{N},\ u_{n+1} = f(u_n)$, avec $f(x) = \ln(1+x)$.

– L'étude de la fonction $x \mapsto f(x) - x$ montre :

$$\forall x > -1,\ \ f(x) \leqslant x,\ \text{ et } f(x) = x \Leftrightarrow x = 0$$

– On montre par récurrence :

$$\forall n \in \mathbf{N},\quad u_n \text{ existe et } u_n \geqslant 0.$$

– $\forall x \geqslant 0,\ f(x) \leqslant x\,;\ \forall n \in \mathbf{N}, u_n \geqslant 0$, donc :

$$\forall n \in \mathbf{N},\quad u_{n+1} = f(u_n) \leqslant u_n,$$

donc (u_n) est décroissante.

– La suite (u_n) est décroissante et minorée par 0, elle est donc convergente. Sa limite ℓ vérifie $\ell \geqslant 0$, et ℓ est un point fixe de f, car f est continue sur l'intervalle $[0, +\infty[$ (attention à considérer l'intervalle **fermé**). Or $f(x) = x \Leftrightarrow x = 0$, donc $\ell = 0$.

• $u_0 = a \geqslant 0\,;\ \forall n \in \mathbf{N},\ u_{n+1} = f(u_n)$, avec $f(x) = x\ln(1+x)$.

– On établit par récurrence : $\forall n \in \mathbf{N},\ u_n \geqslant 0$.

– f est continue sur $[0, +\infty[$, et les points fixes de f sont 0 et $\mathrm{e} - 1$. Donc, **si** (u_n) converge, alors sa limite est 0 ou $\mathrm{e} - 1$.

– f est croissante sur $[0, +\infty[$ et $\forall n \in \mathbf{N}, u_n \geqslant 0$, donc (u_n) est monotone.

– Si $u_0 > \mathrm{e} - 1$, alors $\ln(1 + u_0) > \ln(1 + \mathrm{e} - 1) = 1$, par conséquent $u_1 = u_0 \ln(1 + u_0) > u_0$. La suite (u_n) est donc croissante.

On a donc, pour tout $n \in \mathbf{N}, u_n \geqslant u_0 > \mathrm{e} - 1$: (u_n) ne peut pas converger ni vers $\mathrm{e} - 1$, ni vers 0, elle est donc divergente, et puisqu'elle est croissante : $\lim\limits_{n \to +\infty} u_n = +\infty$.

– Si $u_0 = \mathrm{e} - 1$, alors, par récurrence : $\forall n \in \mathbf{N},\ u_n = \mathrm{e} - 1$.

– Si $0 < u_0 < \mathrm{e} - 1$, alors $u_1 = u_0 \ln(1 + u_0) < u_0$. La suite (u_n) est donc décroissante, et minorée par 0, donc convergente. (u_n) ne peut converger vers $\mathrm{e} - 1$, puisque pour tout n, $u_n \leqslant u_0 < \mathrm{e} - 1$. (u_n) converge donc vers 0.

– Si $u_0 = 0$, alors, par récurrence : $\forall n \in \mathbf{N},\ u_n = 0$.

• (Utilisation pour obtenir une valeur approchée d'une solution à une équation)

– L'étude de la fonction g définie par $g(x) = 2 - x - 2\mathrm{e}^{-x}$ fournit :

x	$-\infty$		$\ln 2$		$+\infty$
$f(x)$	$-\infty$	$\nearrow$	$1 - \ln 2$	$\searrow$	$-\infty$

Le théorème de la bijection monotone montre alors que l'équation $g(x) = 0$ admet deux solutions. Une d'elles est égale à 0, l'autre, r, est supérieure à $\ln 2$. On a $g(1) > 0$, $g(2) < 0$, et g est strictement décroissante sur $[\ln, +\infty[$, donc $1 < r < 2$.

– Soit (u_n) définie par

$$u_0 = 1\,; \quad \forall n \in \mathbf{N}, \quad u_{n+1} = f(u_n) \text{ avec } f(x) = 2\left(1 - \mathrm{e}^{-x}\right)$$

f est continue sur $\mathbf{R}$, donc **si** (u_n) converge, alors sa limite ℓ vérifie l'équation $f(x) = x$; or cette équation est équivalente à $g(x) = 0$, donc $\ell = 0$ ou $\ell = r$.

f est croissante sur $\mathbf{R}$, donc (u_n) est monotone.

$u_1 \geqslant u_0$, donc (u_n) est croissante.

Par récurrence : $\forall n \in \mathbf{N}$, $u_n \leqslant r$.

(u_n) est croissante est majorée, elle est donc convergente, et sa limite ℓ est égale à r, puisque avec $u_0 = 1$ et (u_n) croissante, elle ne peut être égale à 0. u_n (dont on peut programmer le calcul) constitue une valeur approchée de r qu'on ne connaît pas.

3.3 Utilisation de la formule des accroissements finis

On rappelle la formule des accroissements, sous ses deux versions (voir § 1.3.2) :

Première version : Si $m \leqslant f' \leqslant M$ sur $[a, b]$, alors

$$m(b-a) \leqslant f(b) - f(a) \leqslant M(b-a)$$

Deuxième version : Si $|f'| \leqslant k$ sur l'intervalle I, alors,

$$\forall x_1, x_2 \in I, \quad |f(x_1) - f(x_2)| \leqslant k|x_1 - x_2|$$

Pour montrer que la suite (u_n) converge vers ℓ en utilisant la formule des accroissements finis, la démarche en général suivie par l'énoncé est la suivante :

- On établit

– Pour tout $n \in \mathbf{N}$, u_n appartient à l'intervalle I.

– Il existe ℓ appartenant à I tel que $f(\ell) = \ell$.

– $|f'| \leqslant k < 1$ sur I.

- La formule des accroissements finis fournit alors :

$$\forall n \in \mathbf{N}, \quad |f(u_n) - f(\ell)| \leqslant k|u_n - \ell|, \text{ c'est-à-dire}$$
$$\forall n \in \mathbf{N}, \quad |u_{n+1} - \ell| \leqslant k|u_n - \ell|$$

- On montre alors par récurrence :

$$\forall n \in \mathbf{N}, \quad |u_n - \ell| \leqslant k^n\,|u_0 - \ell|$$

En effet : $|u_0 - \ell| \leqslant k^0 |u_0 - \ell|$ car $k^0 = 1$,
et si $|u_n - \ell| \leqslant k^n |u_0 - \ell|$ pour quelque $n \in \mathbf{N}$, alors

$$|u_{n+1} - \ell| \leqslant k\,|u_n - \ell| \leqslant kk^n\,|u_0 - \ell| = k^{n+1}\,|u_0 - \ell|$$

D'où la conclusion.

• $0 < k < 1$, donc $\lim\limits_{n \to +\infty} k^n = 0$, donc par passage à la limite :

$$\lim_{n \to +\infty} u_n = \ell$$

Vous répondrez bien sûr aux questions de l'énoncé, mais la démarche est à retenir, les questions pouvant être plus ou moins détaillées.

Exemple utilisant la première version de la formule des accroissements finis : il s'agit de l'algorithme de Héron l'Ancien, pour la recherche de valeurs approchées de $\sqrt{x}$, avec ici $x = 3$.

$u_0 = a > 0\,;\ \forall n \in \mathbf{N},\ u_{n+1} = f(u_n)$, avec $f(x) = \frac{1}{2}\left(x + \frac{3}{x}\right)$

– L'étude de f montre que le minimum de f sur $]0, +\infty[$ est atteint en $\sqrt{3}$, et égal à $\sqrt{3}$ ($\sqrt{3}$ est donc un point fixe de f).

– On montre par récurrence : $\forall n \in \mathbf{N}^*,\ u_n \geqslant \sqrt{3}$.

– Pour tout $x \in \left[\sqrt{3}, +\infty\right[$:

$$0 \leqslant f'(x) = \frac{1}{2}\left(1 - \frac{3}{x^2}\right) \leqslant \frac{1}{2}$$

– $\sqrt{3} \leqslant u_n$ et $0 \leqslant f' \leqslant \frac{1}{2}$ sur $\left[\sqrt{3}, u_n\right]$, la formule des accroissements finis donne alors :

$$\forall n \in \mathbf{N}^*,\quad 0 \leqslant f(u_n) - f\left(\sqrt{3}\right) \leqslant \frac{1}{2}\left(u_n - \sqrt{3}\right)\text{, donc}$$

$$\forall n \in \mathbf{N}^*,\quad 0 \leqslant u_{n+1} - \sqrt{3} \leqslant \frac{1}{2}\left(u_n - \sqrt{3}\right)$$

– On obtient par récurrence :

$$\forall n \in \mathbf{N}^*,\quad 0 \leqslant u_n - \sqrt{3} \leqslant \left(\frac{1}{2}\right)^{n-1}\left(u_1 - \sqrt{3}\right)$$

– Et comme $\lim\limits_{n \to +\infty} \left(\frac{1}{2}\right)^{n-1} = 0$:

$$\lim_{n \to +\infty} u_n = \sqrt{3}$$

En prenant $u_0 = a = 1$, on obtient

$$\forall n \in \mathbf{N}^*, \quad 0 \leqslant u_n - \sqrt{3} \leqslant \left(\frac{1}{2}\right)^{n-1} \left(2 - \sqrt{3}\right) \leqslant \left(\frac{1}{2}\right)^{n-2}$$

u_n est donc une valeur approchée par excès de $\sqrt{3}$ à moins de $0,5^{n-2}$ près. La variable `u` va recevoir les valeurs successives $u_1, u_2, u_3, \dots$ et la variable `puiss` les valeurs successives 2 ; 1 ; 0,5 ; ... jusqu'à ce que son contenu soit inférieur à 10^{-3}, par exemple. On affiche alors `u`, qui est une valeur approchée de $\sqrt{3}$ à moins de 10^{-3} près :

```
 rogram heron;
var u, uiss:real;n:integer;
function f(x :real) :real ; begin
f :=0.5*(x+3 x) ; end ;
BEGIN
u=2;n:=1; uiss:=2; {n donnera le rang du terme en cours
de calcul}
re eat
u:=f(u);n:=n+1; uiss:= uiss 2;
until  uiss<1.e-3;
writeln(n,u); {le terme qui convient est affiché avec son rang}
END.
```

4. Séries numériques

4.1 Définitions

Soit $(u_n)_{n\in\mathbf{N}}$ une suite numérique. La **série de terme général** $u_n, n \in \mathbf{N}$, est la suite des **sommes partielles** $(S_n)_{n\in\mathbf{N}}$, avec

$$S_n = \sum_{k=0}^{n} u_k$$

La série est dite **convergente** ssi la suite $(S_n)_{n\in\mathbf{N}}$ est convergente ; la limite S de (S_n) est alors appelée la **somme** de la série, et on note :

$$S = \lim_{n\to+\infty} \sum_{k=0}^{n} u_k = \sum_{k=0}^{+\infty} u_k$$

On adaptera sans peine ces définitions au cas où la suite (u_n) est définie sur $\mathbf{N}^*$, par exemple.

On abrégera « série de terme général u_n, $n \in \mathbf{N}$ » en « série $\sum_{n\in\mathbf{N}} u_n$ », sans préjuger de la convergence de la série, et on veillera à ne pas

confondre cette notation avec la notation $\sum_{n=0}^{+\infty} u_n$, qui est réservée aux séries convergentes. (Par contre on a $\sum_{n=0}^{+\infty} u_n = \sum_{k=0}^{+\infty} u_k$, n ayant le statut de variable muette.)

On ne change pas la nature d'une série si on en change un nombre fini de termes, mais la valeur de la somme peut changer.

4.2 Séries usuelles

Séries géométriques et séries géométriques dérivées

géométrique

- La série de terme général x^n, $n \in \mathbf{N}$, converge ssi $|x| < 1$, et

$$\forall x \in \left]-1, 1\right[, \quad \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$$

- La série de terme général nx^{n-1}, $n \in \mathbf{N}^*$, converge ssi $|x| < 1$, et

$$\forall x \in \left]-1, 1\right[, \quad \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$

- La série de terme général $n(n-1)x^{n-2}$, $n \in \mathbf{N}^*$, $n \geqslant 2$, converge ssi $|x| < 1$, et

$$\forall x \in \left]-1, 1\right[, \quad \sum_{k=2}^{+\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3}$$

Démonstration. Si $|x| \geqslant 1$, le terme général des séries concernées ne tend pas vers 0, elles sont donc divergentes (voir § 2.4.3). On suppose donc dans la suite $|x| < 1$.

Pour la série $\sum x^n$: soit $S_n(x) = \sum_{k=0}^{n} x^k$. D'après l'identité géométrique, on a :

$$S_n(x) = \frac{1-x^{n+1}}{1-x} \underset{n\to+\infty}{\longrightarrow} \frac{1}{1-x}$$

car $|x| < 1$, donc $\left(x^{n+1}\right)$ tend vers 0. Et le résultat est établi.

Pour la série $\sum nx^{n-1}$:

$$(1-x)\sum_{k=1}^{n} kx^{k-1} = \sum_{k=1}^{n} kx^{k-1} - \sum_{k=1}^{n} kx^k = \sum_{i=0}^{n-1} (i+1)x^i - \sum_{i=1}^{n} ix^i$$

(Changement d'indice $i = k - 1$ dans la première somme, $i = k$ dans la deuxième.) En regroupant les termes en x^i quand c'est possible :

$$(1-x)\sum_{k=1}^{n} kx^{k-1} = 1 + \sum_{i=1}^{n-1} (i+1-i)\, x^i - nx^n = S_{n-1}(x) - nx^n$$

D'où la conclusion par passage à la limite, car (nx^n) tend vers 0.

Pour la série $\sum n(n-1)\, x^{n-2}$, on procède de même, en multipliant la somme partielle par $1 - x$.

• Vous retiendrez ces trois formules, utiles entre autres en probabilités dans l'étude de la loi géométrique, voir § 8.3.1.

Vous justifierez leur usage en invoquant la série géométrique, resp. la série géométrique dérivée première, resp. la série géométrique dérivée seconde, de raison $x \in \left]-1, 1\right[$.

• Pour la mémorisation, remarquez : ($\frac{\mathrm{d}}{\mathrm{d}x}$ dérivation par rapport à x)

$$\frac{\mathrm{d}}{\mathrm{d}x} x^k = kx^{k-1} \quad \text{et} \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right) = \frac{1}{(1-x)^2}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} kx^{k-1} = k(k-1)\, x^{k-2} \quad \text{et} \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{(1-x)^2}\right) = \frac{2}{(1-x)^3}$$

• Pour la démonstration de la deuxième formule, vous pouvez retenir ce calcul très classique et souvent demandé :

Soit $f_n(x) = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$; f_n est dérivable sur $\left]-1, 1\right[$, et

$$f_n{}'(x) = \sum_{k=1}^{n} kx^{k-1} = \frac{1-(n+1)\,x^n + nx^{n+1}}{(1-x)^2} \underset{n\to+\infty}{\longrightarrow} \frac{1}{(1-x)^2}$$

Remarque. On a la généralisation suivante, que l'on établit par récurrence :

$$\forall x \in \left]-1, 1\right[, \forall n \in \mathbf{N}^*, \quad \sum_{k=n}^{+\infty} k(k-1)\dots(k-n+1)\, x^{k-n} = \frac{n!}{(1-x)^{n+1}}$$

Séries de Riemann

Soit $\alpha \in \mathbf{R}$. La série de terme général $\frac{1}{n^\alpha}$, $n \in \mathbf{N}^*$, est convergente ssi $\alpha > 1$.

En particulier, la série $\sum_{n\in\mathbf{N}^*} \frac{1}{n}$ (série harmonique) est divergente, car $\lim\limits_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k} = +\infty$ comme on l'a montré au § 2.1.3.

Si $\alpha \leqslant 0$, la série diverge car son terme ne général ne tend pas vers 0, voir § 2.4.3.

Si $\alpha > 0$, on établit le résultat par comparaison avec une intégrale généralisée, voir § 3.4.

Les séries de Riemann servent très souvent de séries de référence dans les critères de convergence, voir § 2.4.3.

Série exponentielle

Pour tout $x \in \mathbf{R}$, la série de terme général $\frac{x^n}{n!}$, $n \in \mathbf{N}$, est convergente, de somme e^x :

$$\forall x \in \mathbf{R}, \quad \sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x$$

En particulier, $1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots = e$.

On peut démontrer la formule générale en utilisant l'inégalité de Taylor-Lagrange, voir § 3.2.5.

On utilise ce résultat en probabilités, dans l'étude de la loi de Poisson, voir § 8.3.2.

4.3 Critères de convergence

Un critère négatif

Si la série $\sum_{n\in\mathbf{N}} u_n$ converge, alors $\lim\limits_{n\to+\infty} u_n = 0$.

En effet, si $\sum u_n$ converge alors par définition $\lim\limits_{n\to+\infty} S_n = S \in \mathbf{R}$, avec $S_n = \sum_{k=0}^{n} u_k$; donc $u_n = S_n - S_{n-1}$ tend vers 0 quand n tend vers $+\infty$.

Il s'agit d'un critère négatif, il sert le plus souvent à montrer qu'une série n'est pas convergente. On l'a utilisé plusieurs fois dans le § 4.2.

Attention, si le terme général tend vers 0, on ne peut rien dire : la série de terme général $\frac{1}{n^2}$ converge, la série de terme général $\frac{1}{n}$ diverge.

Séries à termes positifs

- $\forall n \geqslant n_0, 0 \leqslant u_n \leqslant v_n$; $\sum v_n$ converge $\Rightarrow \sum u_n$ converge.
$\forall n \geqslant n_0, 0 \leqslant u_n \leqslant v_n$; $\sum u_n$ diverge $\Rightarrow \sum v_n$ diverge.
- $\forall n \geqslant n_0, u_n \geqslant 0, v_n \geqslant 0$; $u_n \underset{n\to+\infty}{\sim} v_n \Rightarrow$
$\sum u_n$ et $\sum v_n$ sont de même nature
(simultanément convergentes ou divergentes).
- $\forall n \geqslant n_0,\ u_n \geqslant 0, v_n \geqslant 0$; $u_n \underset{n\to+\infty}{=} \circ(v_n)$;
$\sum v_n$ converge $\Rightarrow \sum u_n$converge.

Pour établir le premier critère, on montre que la suite des sommes partielles $S_n = \sum_{n=n_0}^{n} u_k$ est croissante, car $u_k \geqslant 0$, et majorée, car $S_n = \sum_{k=n_0}^{n} u_k \leqslant \sum_{k=n_0}^{n} v_k \leqslant \sum_{k=n_0}^{+\infty} v_k \in \mathbf{R}$.

Ces critères tombent en défaut pour les séries qui ne sont pas à termes positifs. Vous signalerez donc que la série est à termes positifs, même si cela est évident.

La série de terme général $\frac{e^{\frac{1}{n}}}{n^2}$, $n \geqslant 1$, est convergente, car

$$\forall n \geqslant 1,\ \frac{e^{\frac{1}{n}}}{n^2} \geqslant 0,\ \frac{e^{\frac{1}{n}}}{n^2} \underset{n\to+\infty}{\sim} \frac{1}{n^2},\ \text{et} \sum_{n\geqslant 1} \frac{1}{n^2} \text{ converge (série de Riemann)}.$$

Séries à termes quelconques

La série de terme général u_n est dite **absolument convergente** ssi la série de terme général $|u_n|$ est convergente.

Théorème. Une série absolument convergente est convergente.

Théorème admis. On peut préciser : si la série $\sum_{n\geqslant n_0} u_n$ est absolument convergente, alors elle est convergente, et

$$\left|\sum_{n=n_0}^{+\infty} u_n\right| \leqslant \sum_{n=n_0}^{+\infty} |u_n|$$

Attention, la réciproque de ce théorème est fausse, une série peut être convergente sans être absolument convergente.

4.4 Calculs de sommes de séries

Théorème. Soit α, β deux nombres réels, et soit deux séries convergentes, de termes généraux $u_n, v_n, n \geqslant n_0$. Alors la série de terme général $\alpha u_n + \beta v_n$, $n \geqslant n_0$, est convergente, et on a :

$$\sum_{n=n_0}^{+\infty} (\alpha u_n + \beta v_n) = \alpha \sum_{n=n_0}^{+\infty} u_n + \beta \sum_{n=n_0}^{+\infty} v_n$$

Une combinaison linéaire (voir chapitre 5) de séries convergentes est donc une série convergente.

Si la série $\sum u_n$ diverge, alors la série $\sum \alpha u_n$ diverge (avec $\alpha \neq 0$), mais attention, la somme de deux séries divergentes peut être une série convergente.

• Soit $x \in]-1, 1[$. Montrons que la série $\sum_{n\geqslant 1} nx^n$ est convergente, et calculons sa somme. Sous réserve de convergence, on a :

$$\sum_{n=1}^{+\infty} nx^n = x \sum_{n=1}^{+\infty} nx^{n-1}$$

Or la série de terme général nx^{n-1}, $n \in \mathbf{N}^*$, est convergente, de somme $\frac{1}{(1-x)^2}$, car c'est une série géométrique dérivée de raison $x \in]-1, 1[$. Donc la série de terme général nx^{n-1}, $n \in \mathbf{N}^*$, est convergente, et on a :

$$\sum_{n=1}^{+\infty} nx^n = \frac{x}{(1-x)^2}$$

• Montrons que la série $\sum_{n\geqslant 1} n^2x^n$ est convergente et calculons sa somme. On peut montrer qu'il s'agit d'une série convergente en utilisant les critères de convergence des séries à terme positifs :

$$\forall n \geqslant 1,\ \ n^2x^n \geqslant 0\,;\ \ n^2x^n = \circ\left(\frac{1}{n^2}\right) \text{ car } \lim_{n\to+\infty} n^4x^n = 0 \text{ car } |x| < 1\,;$$

$\sum_{n\geqslant 2} \frac{1}{n^2}$ est convergente (série de Riemann) ; donc la série $\sum_{n\geqslant 1} n^2x^n$ est convergente, mais cela ne permet pas de calculer sa somme. Pour ce

calcul, on écrit, sous réserve de convergence,

$$\sum_{n=1}^{+\infty} n^2 x^n = \sum_{n=1}^{+\infty} (n(n-1)+n)\, x^n$$
$$= x^2 \sum_{n=2}^{+\infty} n(n-1)\, x^{n-2} + x \sum_{n=1}^{+\infty} n x^{n-1}$$

On a donc une combinaison linéaire de séries convergentes (séries géométriques dérivées premières et secondes, de raison $x \in]-1, 1[$), dont les sommes sont connues ; la série est donc convergente, et tous calculs faits on trouve

$$\sum_{n=1}^{+\infty} n^2 x^n = \frac{x(x+1)}{(1-x)^3}$$

On peut aussi obtenir la somme d'une série quand les sommes partielles se prêtent à des simplifications (« dominos », « téléscopage »).

Montrons que la série $\sum_{k\geqslant 2} \frac{1}{k(k-1)}$ est convergente, et calculons sa somme, en revenant à la définition. La somme partielle de rang n est :

$$S_n = \sum_{k=2}^{n} \frac{1}{k(k-1)} = \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k}\right)$$
$$= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) = 1 - \frac{1}{n}$$

Par conséquent la série est convergente, et

$$\sum_{k=2}^{+\infty} \frac{1}{k(k-1)} = 1$$

5. Suites définies implicitement

On entend par là les suites (u_n) où u_n est défini par une condition du type $f_n(u_n) = 0$. Il s'agit d'un thème d'exercices, aucune connaissance spécifique n'est nécessaire. Les outils utilisés sont :

• le théorème de la bijection monotone, pour prouver l'existence et l'unicité de u_n ;

• le sens de variations de f_n, pour majorer (ou minorer) u_n. Par exemple si f_n est strictement croissante, $f_n(0) < 0$ et $f_n(1) > 0$, alors $0 < u_n < 1$;

• le sens de variations de f_n peut permettre aussi d'étudier le sens de variations de (u_n). En effet, supposons par exemple que f_n est strictement croissante pour tout n, et que $f_{n+1}(u_n) > 0$. On a alors : $f_{n+1}(u_{n+1}) = 0$; $f_{n+1}(u_n) > 0$; f_{n+1} strictement croissante ; donc $u_n > u_{n+1}$, et (u_n) est strictement décroissante. Un tableau de variation peut aider l'intuition :

x	u_{n+1}		u_n
$f_{n+1}(x)$	0	$\nearrow$	> 0

• la définition même de u_n par la condition $f_n(u_n) = 0$. On ne cherchera pas à déterminer une expression explicite de u_n, moins que cela ne soit demandé par l'énoncé.

– Pour $n \in \mathbf{N}^*$, le théorème de la bijection monotone appliqué à $f_n : x \mapsto x^5 + nx - 1$ montre qu'il existe un unique réel u_n tel que $f_n(u_n) = 0$.

– $f_n(0) = -1 < 0$, $f_n\left(\frac{1}{n}\right) = \frac{1}{n^5} > 0$, f_n est strictement croissante, donc $0 < u_n < \frac{1}{n}$, et par encadrement : (u_n) converge vers 0.

– Pour tout $n \geqslant 1$, $f_{n+1}(u_n) = u_n > 0$, $f_{n+1}(u_{n+1}) = 0$, f_{n+1} est strictement croissante, donc $u_{n+1} < u_n$, la suite (u_n) est décroissante.

– $f_n(u_n) = u_n^5 + nu_n - 1 = 0$, donc $nu_n = 1 - u_n^5$ tend vers 1 quand n tend vers $+\infty$, donc u_n est équivalent à $\frac{1}{n}$ quand n tend vers $+\infty$.

– À partir de $u_n^5 + nu_n - 1 = 0$, on peut alors obtenir un équivalent de $\frac{1}{n} - u_n$ quand n tend vers $+\infty$:

$$\frac{u_n^5}{n} + u_n - \frac{1}{n} = 0 \,; \qquad \frac{1}{n} - u_n = \frac{u_n^5}{n} \underset{n\to+\infty}{\sim} \frac{1}{n^6}$$



Calcul intégral

3

1. Primitives

Définition. Soit f une fonction définie sur l'intervalle I. On dit que $F : I \to \mathbf{R}$ est une **primitive** de f sur I ssi

$$\forall x \in I, F'(x) = f(x)$$

Propriétés

• Si f est continue sur l'intervalle I, alors f admet des primitives sur I.

• Soit F et G deux primitives de f sur I. Alors il existe $k \in \mathbf{R}$ tel que

$$\forall x \in I, \quad G(x) = F(x) + k$$

• Soit f continue sur l'intervalle I, $x_0 \in I$, $y_0 \in \mathbf{R}$. Alors il existe une unique primitive F de f sur I tel que $F(x_0) = y_0$.

La première propriété est admise. On démontre la deuxième propriété à l'aide de l'inégalité des accroissements finis appliquée à $H = F - G$ sur I. La troisième propriété se déduit de la deuxième.

Calcul des primitives

a) Primitives des fonctions usuelles

fonction	primitive	commentaire
$x \mapsto a$	$x \mapsto ax$	sur $\mathbf{R}$
$x \mapsto x^r$	$x \mapsto \dfrac{x^{r+1}}{r+1}$	sur tout intervalle où la fonction $x \mapsto x^r$ est continue, et donc sur : $\mathbf{R}$ si r est un entier naturel $\mathbf{R}^{+*}$ et $\mathbf{R}^{-*}$ si r est un entier relatif < -1 $\mathbf{R}^{+}$ si r est un réel positif non entier $\mathbf{R}^{+*}$ si r est un réel négatif non entier
$x \mapsto \dfrac{1}{x}$	$x \mapsto \ln\lvert x\rvert$	sur $\mathbf{R}^{+*}$ et $\mathbf{R}^{-*}$
$x \mapsto e^x$	$x \mapsto e^x$	sur $\mathbf{R}$

Exemples :

– Sur $\mathbf{R}$, une primitive de $x \mapsto x$ est $x \mapsto \dfrac{x^2}{2}$

– Sur $\mathbf{R}^{+*}$ et $\mathbf{R}^{-*}$, une primitive de $x \mapsto \dfrac{1}{x^2}$ est $x \mapsto -\dfrac{1}{x}$

– Sur $\mathbf{R}^{+*}$, une primitive de $x \mapsto \dfrac{1}{\sqrt{x}}$ est $x \mapsto 2\sqrt{x}$

– Sur $\mathbf{R}^{+}$, une primitive de $x \mapsto \sqrt{x} = x^{\frac{1}{2}}$ est $x \mapsto \dfrac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1} = \dfrac{2}{3}x\sqrt{x}$

Vous devriez retenir sans peine les trois premières formules (d'usage très fréquent), car vous avez souvent dérivé les fonctions x^2, $\frac{1}{x}$, $\sqrt{x}$!

b) Règles de calcul

fonction	primitive	commentaire
u	U	u continue sur l'intervalle I
v	V	v continue sur I
$au + bv$	$aU + bV$	sur I, avec a, b constantes I
$u'f(u)$	$F(u)$	sur I, avec $u \in C^1(I)$ et f continue sur $u(I)$, de primitive F

Exemples d'utilisation de la dernière formule. Sur tout intervalle où la fonction « à intégrer » (c'est-à-dire dont on cherche une primitive) est continue :

- Une primitive de $u'u^r$ est $\dfrac{u^{r+1}}{r+1}(r \in \mathbf{R} \setminus \{-1\})$.

En particulier :

– Une primitive de $u'u$ est $\dfrac{u^2}{2}$.

– Une primitive de $\dfrac{u'}{u^2}$ est $-\dfrac{1}{u}$.

– Une primitive de $\dfrac{u'}{\sqrt{u}}$ est $2\sqrt{u}$.

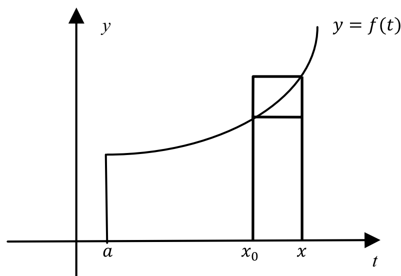
- Une primitive de $\dfrac{u'}{u}$ est $\ln|u|$.
- Une primitive de $u'\mathrm{e}^u$ est e^u.

Ici aussi formules d'emploi très fréquent, à mettre en parallèle respectivement avec les primitives des fonctions x^r, x, $\frac{1}{x^2}$, $\frac{1}{\sqrt{x}}$, $\frac{1}{x}$, e^x.

2. Intégrale définie

2.1 Définition, interprétation graphique

Soit : f une fonction continue, positive et croissante sur l'intervalle I ; $a, x, x_0 \in I$ tels que $a \leqslant x_0 < x$; $F(x)$ l'aire du domaine plan limité par l'axe des abscisses, les droites d'équation $t = a$, $t = x$ et la courbe représentative de f.



On voit sur le dessin, en se souvenant de la formule qui donne l'aire d'un rectangle, que

$$f(x_0)(x - x_0) \leqslant F(x) - F(x_0) \leqslant f(x)(x - x_0)$$

Par conséquent

$$f(x_0) \leqslant \frac{F(x) - F(x_0)}{x - x_0} \leqslant f(x)$$

f est continue en x_0, donc en passant à la limite dans cette double inégalité, on obtient $F'_d(x_0) = f(x_0)$. On obtient de la même manière $F'_g(x_0) = f(x_0)$, donc $F'(x_0) = f(x_0)$ en tout $x_0 \in I$. F est donc la primitive de f sur $[a, b]$ qui s'annule en a. On est conduit ainsi à donner la définition suivante :

Définition. Soit I un intervalle, a et b deux éléments de I, et f une fonction continue sur I. L'**intégrale de** a **à** b **de la fonction** f est le nombre réel défini par :

$$\int_a^b f(t)\, \mathrm{d}t = F(b) - F(a)$$

où F est une primitive de f sur I.

• f est continue sur I, donc admet des primitives. Deux primitives de f sur I diffèrent d'une constante, la valeur de l'intégrale ne dépend donc pas de la primitive choisie.

• On lit « somme de a à b de $f(t)\mathrm{d}t$ ». Ne pas omettre le symbole différentiel « $\mathrm{d}t$ », qui indique la variable d'intégration t.

• Dans l'écriture $\int_a^b f(t)\, \mathrm{d}t$, t est une variable muette, a et b sont des variables libres. La valeur de l'intégrale ne dépend pas de t, mais dépend *a priori* de a et b. Ainsi

$$\int_a^b f(t)\, \mathrm{d}t = \int_a^b f(x)\, \mathrm{d}x$$

Interprétation en termes d'aire. Soit f une fonction continue et positive ou nulle sur l'intervalle $[a, b]$.

Le plan est rapporté au repère orthogonal $\left(O; \vec{i}, \vec{j}\right)$.

Alors l'intégrale $\int_a^b f(t)\, \mathrm{d}t$ est égale à l'aire du domaine plan limité par l'axe des abscisses (Ox), les droites d'équation $y = a$, $y = b$, et la courbe d'équation $y = f(x)$. Cette aire est exprimée en unités d'aire (u.a), aire du rectangle de cotés $\|\vec{i}\|$, $\|\vec{j}\|$.

Attention aux hypothèses $a \leqslant b$ et $f \geqslant 0$. Si elles ne sont pas vérifiées, le résultat ne subsiste pas, voir § 3.3.4.

2.2 Calculs d'intégrales

Intégration « à vue »

On écrit

$$\int_a^b f(t)\,\mathrm{d}t = [F(t)]_a^b = F(b) - F(a)$$

où F est une primitive de f sur I , avec a et b dans I.

- $\displaystyle\int_0^1 \mathrm{e}^{-t}\,\mathrm{d}t = \left[-\mathrm{e}^{-t}\right]_0^1 = -\mathrm{e}^{-1} - \left(-\mathrm{e}^0\right) = 1 - \frac{1}{\mathrm{e}}$

Une primitive de $u'\mathrm{e}^u$ est e^u. Ici $u = -t$, $u' = -1$, $\mathrm{e}^{-t} = -u'\mathrm{e}^u$, d'où le résultat.

- $\displaystyle\int_0^1 \frac{x}{x^2+1}\,\mathrm{d}x = \left[\frac{1}{2}\ln\left(x^2+1\right)\right]_0^1 = \frac{1}{2}\ln 2$

On repère que la fonction à intégrer est de la forme $k\frac{u'}{u}$, dont une primitive est $k\ln|u|$. Ici $u = x^2+1 > 0$, $u' = 2x$, $k = \frac{1}{2}$, d'où le résultat.

- Même principe pour

$$\int_0^1 \frac{x}{\left(x^2+1\right)^2}\,\mathrm{d}x = \left[-\frac{1}{2}\frac{1}{x^2+1}\right]_0^1 = \frac{1}{4}$$

Une primitive de $\frac{u'}{u^2}$ est $-\frac{1}{u}$. On commence par écrire $\frac{1}{x^2+1}$ dans le crochet, on dérive mentalement, et on ajuste la constante multiplicative.

- $\displaystyle\int_0^1 \frac{x}{\sqrt{x^2+1}}\,\mathrm{d}x = \left[\sqrt{x^2+1}\right]_0^1 = 1$
- $\displaystyle\int_0^1 \left(\frac{t}{2} + \frac{1}{2t+1} + 2\mathrm{e}^{3t} - 1\right)\mathrm{d}t = \left[\frac{t^2}{4} + \frac{1}{2}\ln(2t+1) + \frac{2}{3}\mathrm{e}^{3t} - t\right]_0^1$

On trouve une primitive d'une somme en prenant une somme de primitives, une primitive de au (a constante) en multipliant une primitive de u par a. Tous calculs fait on trouve $\frac{\ln 3}{3} + \frac{2\mathrm{e}^3}{3} - \frac{13}{12}$.

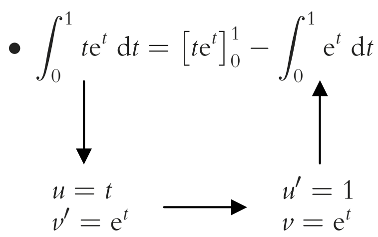
Intégration par parties

Soit I un intervalle, a et b deux éléments de I, u et v deux fonctions de classe C^1 sur l'intervalle I.

Alors

$$\int_a^b u(t)v'(t)\,\mathrm{d}t = [u(t)v(t)]_a^b - \int_a^b u'(t)v(t)\,\mathrm{d}t$$

Disposition pratique :



Sous l'intégrale de départ, on écrit u et v' tels que le produit uv' soit la fonction sous le signe somme. On calcule u' et v dont le produit donne la fonction sous l'intégrale d'arrivée.

Il reste à calculer l'intégrale d'arrivée :

$$I = e - \left[e^t\right]_0^1 = e - (e - 1) = 1$$

- N'oubliez pas de mentionner – et de vérifier – que les fonctions u et v sont de classe C^1 sur un intervalle I auquel appartiennent les bornes a, b ($I = [0\,;\,1]$ convient dans l'exemple précédent).
- L'intégrale d'arrivée doit être plus simple que l'intégrale de départ. Si ce n'est pas le cas, faites un autre choix pour u et v'.
- Avant de commencer une intégration par parties, vérifiez si une autre méthode plus simple ne s'applique pas (intégration à vue).

Changement de variable

Soit I un intervalle, a et b deux éléments de I, u une fonction de classe C^1 sur l'intervalle I, f une fonction continue sur l'intervalle $u(I)$.

Alors

$$\int_a^b u'(t) f(u(t)) \, dt = \int_{u(a)}^{u(b)} f(y) \, dy$$

Grâce à la notation différentielle de la dérivée ($\frac{dy}{dt}$ pour $y'(t)$), on utilise facilement cette formule, qu'il est inutile de retenir telle quelle.

$$I = \int_0^1 \frac{1 + \sqrt{x+1}}{x+1} \, dx$$

Soit $y = \sqrt{x+1}$; $\dfrac{dy}{dx} = \dfrac{1}{2\sqrt{x+1}} = \dfrac{1}{2y}$, donc $dx = 2y \, dy$, puis

$I = \int_1^{\sqrt{2}} \frac{1+y}{y^2} 2y dy = \int_1^{\sqrt{2}} \left(\frac{2}{y} + 2\right) dy = [2\ln y + 2y]_1^{\sqrt{2}}$
$I = \ln 2 + 2\sqrt{2} - 2$

Application aux intégrales de fonctions paires, impaires

Soit $a > 0$ et f une fonction continue sur $[-a, a]$.

- Si f est paire sur $[-a, a]$, alors

$$\int_{-a}^{a} f(t)\,\mathrm{d}t = 2\int_0^a f(t)\,\mathrm{d}t$$

- Si f est impaire sur $[-a, a]$, alors

$$\int_{-a}^{a} f(t)\,\mathrm{d}t = 0$$

En effet, d'après la relation de Chasles (voir § 3.2.3) :

$$\int_{-a}^{a} f(t)\,\mathrm{d}t = \int_{-a}^{0} f(t)\,\mathrm{d}t + \int_0^a f(t)\,\mathrm{d}t$$

Dans la première intégrale, on fait le changement de variable $y = -t$:
$\frac{\mathrm{d}y}{\mathrm{d}t} = -1$, donc $\mathrm{d}t = -\,\mathrm{d}y$, donc

$$\int_{-a}^{0} f(t)\,\mathrm{d}t = \int_a^0 f(-y)(-\,\mathrm{d}y) = -\int_a^0 f(-y)\,\mathrm{d}y = \int_0^a f(-y)\,\mathrm{d}y$$

$f(-y) = f(y)$ si f est paire, $f(-y) = -f(y)$ si f est impaire, ce qui permet de conclure.

Le changement de variable à effectuer devrait vous être indiqué, sauf s'il s'agit d'un changement de variable affine ($y = at + b$). Le résultat sur les intégrales de fonctions paires ou impaires doit être connu, ainsi que sa démonstration.

2.3 Propriétés de l'intégrale

Avec f, g continues sur I ; $a, b, c \in I$; $\alpha, \beta \in \mathbf{R}$:

- **Propriétés élémentaires**

$$\int_a^a f(t)\,\mathrm{d}t = 0 \;;\quad \int_a^b 0\,\mathrm{d}t = 0 \;;\quad \int_a^b C\,\mathrm{d}t = C(b-a)\ (C \text{ constante})$$

- **Relation de Chasles**

$$\int_a^c f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t + \int_b^c f(t)\,\mathrm{d}t\ ; \quad \int_b^a f(t)\,\mathrm{d}t = -\int_a^b f(t)\,\mathrm{d}t$$

- **Linéarité**

$$\int_a^b (\alpha f(t) + \beta g(t))\,\mathrm{d}t = \alpha \int_a^b f(t)\,\mathrm{d}t + \beta \int_a^b g(t)\,\mathrm{d}t$$

- **Positivité, croissance**

– Si $a \leqslant b$ et $f \geqslant 0$ sur $[a, b]$, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t \geqslant 0$

– Si $a \leqslant b$ et $f \leqslant g$ sur $[a, b]$, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t \leqslant \int_a^b g(t)\,\mathrm{d}t$

- **Intégrale et valeur absolue**

$$\text{Si } a \leqslant b, \text{ alors } \left|\int_a^b f(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}t$$

Tous ces résultats sont des conséquences faciles de la définition de l'intégrale. Par exemple, pour la positivité :

Si $a \leqslant b$ et $f \geqslant 0$ sur $[a, b]$, alors $\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a) \geqslant 0$,

car avec $F' = f \geqslant 0$, F est croissante.

- De la croissance de l'intégrale et de la troisième propriété élémentaire, on déduit le résultat très souvent utilisé :

Si $a \leqslant b$ et $m \leqslant f \leqslant M$ sur $[a, b]$, alors

$$m(b-a) \leqslant \int_a^b f(t)\,\mathrm{d}t \leqslant M(b-a)$$

- De la croissance de l'intégrale et de la relation de Chasles, on déduit, avec f continue sur I, a et b appartenant à I :

– Si $f \geqslant 0$ et $a \leqslant b$, alors $\int_a^b f(t)\,\mathrm{d}t \geqslant 0$;

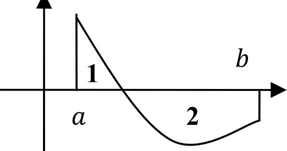
– Si $f \geqslant 0$ et $a \geqslant b$, alors $\int_a^b f(t)\,\mathrm{d}t \leqslant 0$;

– Si $f \leqslant 0$ et $a \leqslant b$, alors $\int_a^b f(t)\,\mathrm{d}t \leqslant 0$;

– Si $f \leqslant 0$ et $a \geqslant b$, alors $\int_a^b f(t)\,\mathrm{d}t \leqslant 0$.

Et dans tous ces cas, la valeur absolue de $\int_a^b f(t)\ \mathrm{d}t$ est égale à l'aire du domaine plan limité par C_f et les droites (Ox), $(x = a)$, $(x = b)$.

Si f change de signe, par exemple comme dans le figure ci-contre, on a $\int_a^b f(t)\ \mathrm{d}t =$ Aire **1** − Aire **2**



• Vous serez amené à utiliser les propriétés de positivité et de croissance de l'intégrale en appliquant le principe :

Pour majorer, minorer, encadrer une intégrale, on majore, minore, encadre la fonction à intégrer, en veillant à ce que les bornes soient « dans le bon sens ». Exemple :

Avec $I_n = \int_0^1 \frac{t^n}{1+t^2}\ \mathrm{d}t$, on a $0 \leqslant I_n \leqslant \frac{1}{n+1}$. En effet :

$\forall t \in [0, 1]\,, 0 \leqslant \frac{t^n}{1+t^2} \leqslant t^n$, donc $0 \leqslant \int_0^1 \frac{t^n}{1+t^2}\ \mathrm{d}t \leqslant \int_0^1 t^n\ \mathrm{d}t = \frac{1}{n+1}$

• Sauf si l'énoncé vous le demande, ne cherchez pas à calculer les valeurs des intégrales qui vous sont proposées. Ce travail est en effet inutile pour utiliser les propriétés ci-dessus.

• **Intégrale fonction de sa borne supérieure**

Théorème. Soit I un intervalle, a un élément de I, f une fonction continue sur I.

Alors l'intégrale

$$\int_a^x f(t)\ \mathrm{d}t$$

existe pour tout $x \in I$ et définit une fonction $x \mapsto \varphi(x) = \int_a^x f(t)\ \mathrm{d}t$ de classe C^1 sur I, de dérivée $x \mapsto f(x)$. Plus précisément, φ est la primitive de f sur I qui s'annule en a.

En effet, l'intégrale existe puisque x et a appartiennent à un même intervalle où f est continue. Les autres propriétés découlent de la définition de l'intégrale.

Avec f continue sur I, l'écriture $\int f(x)\ \mathrm{d}x$ désigne une primitive de f sur I (attention, x n'est plus une variable muette). On peut alors utiliser le théorème ci-dessus pour déterminer une primitive de f en utilisant le calcul intégral (changement de variable, intégration par parties).

L'exemple suivant est très classique :

$$\int \ln x \, \mathrm{d}x = [x \ln x] - \int 1 \, \mathrm{d}x$$

On a effectué une intégration par parties avec

$$u = \ln x\,, \quad v' = 1\ ; \quad u' = \frac{1}{x}\,, \quad v = x$$

u et v sont de classe C^1 sur tout segment de $]0, +\infty[$. Une primitive de la fonction ln sur $]0, +\infty[$ est donc la fonction $x \mapsto x \ln x - x$.

Exemple d'étude d'une fonction définie par une intégrale :

• Pour tout $x \in \mathbf{R}$, l'intégrale $\varphi(x) = \int_x^{2x} \frac{1}{t^4+1} \, \mathrm{d}t$ existe car la fonction $f : t \mapsto \frac{1}{t^4+1}$ est continue sur $\mathbf{R}$. φ est donc une fonction définie sur $\mathbf{R}$.

• Si $x \geqslant 0$, alors $0 \leqslant x \leqslant 2x$; $f \geqslant 0$ sur $[0, +\infty[$, donc $\varphi(x) \geqslant 0$.

Si $x \leqslant 0$, alors $2x \leqslant x \leqslant 0$; $f \geqslant 0$ sur $]-\infty, 0]$, donc $\varphi(x) \leqslant 0$.

$\varphi(0) = 0$.

• Le changement de variable $y = -t$ montre que φ est une fonction impaire ($t = -y,\ \mathrm{d}t = -\,\mathrm{d}y$) :

$\forall x \in \mathbf{R}, \varphi(-x) = \int_{-x}^{-2x} \frac{1}{t^4+1} \, \mathrm{d}t = -\int_x^{2x} \frac{1}{(-y)^4+1} \, \mathrm{d}y = -\varphi(x)$

On étudie donc φ sur $[0, +\infty[$.

• Pour tout $x \in \mathbf{R}$, par définition de l'intégrale,

$\varphi(x) = \int_x^{2x} \frac{1}{t^4+1} \, \mathrm{d}t = F(2x) - F(x)$ avec F primitive de f sur $\mathbf{R}$.

– Ceci permet de préciser le signe de $\varphi(x)$: $F' = f > 0$ donc F est strictement croissante ; si $x > 0$, alors $x < 2x$, $F(x) < F(2x)$, $\varphi(x) > 0$; si $x < 0$, alors $x > 2x$, $F(x) > F(2x)$, $\varphi(x) < 0$.

– Ceci permet aussi d'établir que φ est une fonction dérivable sur $\mathbf{R}$, en tant que somme de deux fonctions dérivables sur $\mathbf{R}$, la fonction $x \mapsto F(2x)$ étant elle-même dérivable en tant que composée de deux fonctions dérivables ; on obtient

$$\forall x \in \mathbf{R}, \quad \varphi'(x) = 2f(2x) - f(x) = \frac{1 - 14x^4}{\left(16x^4 + 1\right)\left(x^4 + 1\right)}$$

φ est donc strictement croissante sur $\left[0, 14^{-\frac{1}{4}}\right]$, strictement décroissante sur $\left[14^{-\frac{1}{4}}, +\infty\right[$.

• Pour la limite en $+\infty$, on encadre l'intégrale $\int_x^{2x} f(t) \, \mathrm{d}t$, et pour cela on encadre la fonction f sur $[x, 2x]$. Or f est décroissante sur $[0, +\infty[$,

donc : $\forall t \in [x, 2x]\,, f(2x) \leqslant f(t) \leqslant f(x)$, puis

$$f(2x)\,(2x - x) \leqslant \int_x^{2x} f(t)\,\mathrm{d}t \leqslant f(x)\,(2x - x)$$

$$\frac{x}{16x^4 + 1} \leqslant \varphi(x) \leqslant \frac{x}{x^4 + 1}$$

et par conséquent, $\lim\limits_{+\infty} \varphi = 0$. φ étant impaire, la limite de φ en $-\infty$ est aussi égale à 0.

2.4 Sommes de Riemann

Très souvent, on ne peut pas trouver la valeur exacte d'une intégrale définie, faute par exemple de pouvoir obtenir une expression explicite d'une primitive de la fonction à intégrer. On doit alors se contenter de chercher des valeurs approchées de l'intégrale. Les sommes de Riemann constituent un premier pas dans cette recherche.

Théorème. Soit f une fonction continue sur l'intervalle $[a, b]$.
Alors

$$\lim_{n \to +\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) = \int_a^b f(t)\,\mathrm{d}t$$

$$\lim_{n \to +\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right) = \int_a^b f(t)\,\mathrm{d}t$$

Ce théorème est admis dans le cas général. Voici les éléments de sa démonstration quand f est de classe C^1 sur $[a, b]$:

- f' est alors continue sur $[a, b]$, donc d'après §1.2.4 , il existe m et M dans $\mathbf{R}$ tel que $m \leqslant f'(t) \leqslant M$ pour tout t dans $[a, b]$;
- Soit $\delta_n = \frac{b-a}{n}$, $a_k = a + k\delta_n$, $a_k \leqslant t \leqslant a_{k+1}$. La formule des accroissements finis fournit :

$$m\,(t - a_k) \leqslant f(t) - f\,(a_k) \leqslant M\,(t - a_k)$$

donc

$$f\,(a_k) + m\,(t - a_k) \leqslant f(t) \leqslant f\,(a_k) + M\,(t - a_k)$$

On intègre sur $[a_k, a_{k+1}]$:

$$\delta_n f\,(a_k) + \frac{m}{2}\delta_n^2 \leqslant \int_{a_k}^{a_{k+1}} f(t)\,\mathrm{d}t \leqslant \delta_n f\,(a_k) + \frac{M}{2}\delta_n^2$$

On additionne ces inégalités, pour k variant de 0 à $n-1$. En posant $S_n = \delta_n \sum_{k=0}^{n-1} f(a_k)$, il vient :

$$S_n + \frac{m}{2}\frac{(b-a)^2}{n} \leqslant \int_a^b f(t)\,\mathrm{d}t \leqslant S_n + \frac{M}{2}\frac{(b-a)^2}{n}$$

$$\frac{m}{2}\frac{(b-a)^2}{n} \leqslant \int_a^b f(t)\,\mathrm{d}t - S_n \leqslant \frac{M}{2}\frac{(b-a)^2}{n}$$

D'où la première formule, en passant à la limite. On procède de même pour la deuxième formule.

On peut facilement définir une fonction PASCAL qui évalue la somme de Riemann $S_n = \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right)$. Cette fonction evalsom aura pour paramètres `a,b` de type réel, et `n` de type entier. La fonction f est définie d'autre part.

```
function evalsom(a,b:real ;n :integer) :real ;
var k:integer ; {variable locale de sommation};
S:real ; {variable locale pour la somme}
begin
  S:=0 ;
  for k:=0 to n-1 do S:=S+f(a+k*(b-a) n);
  evalsom:=S*(b-a) n;
end;
```

2.5 Formules de Taylor

Formule de Taylor avec reste intégral

Théorème. Soit I un intervalle, a un élément de I, $n \in \mathbf{N}$, f une fonction de classe C^{n+1} sur I.

Alors, pour tout x appartenant à I :

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t$$

Démonstration par récurrence : la formule est vraie pour $n = 0$, car

$$\int_a^x f'(t)\,\mathrm{d}t = f(x) - f(a)$$

Supposons la formule vraie pour n fixé dans $\mathbf{N}$, et soit f de classe C^{n+2} sur I. On effectue une intégration par parties sur l'intégrale, avec

$u = f^{(n+1)}$, $v' = \frac{(x-t)^n}{n!}$, $u' = f^{(n+2)}$, $v = -\frac{(x-t)^{n+1}}{(n+1)!}$. u et v sont de classe C^1 sur I, et on obtient la formule à l'ordre $n+1$. Le théorème est ainsi démontré.

Inégalité de Taylor-Lagrange

Théorème. Soit I un intervalle, a un élément de I, $n \in \mathbf{N}$, f une fonction de classe C^{n+1} sur I. Alors, pour tout x appartenant à I :

$$\left| f(x) - \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) \right| \leqslant M \frac{|x-a|^{n+1}}{(n+1)!}$$

Avec M majorant de $\left|f^{(n+1)}\right|$ sur un intervalle contenant a et x.

Il suffit de majorer le reste intégral de la formule précédente, pour obtenir ce résultat. Si $a \leqslant x$:

$$\left| \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, \mathrm{d}t \right| \leqslant \int_a^x \frac{(x-t)^n}{n!} \left|f^{(n+1)}(t)\right| \mathrm{d}t \leqslant M \int_a^x \frac{(x-t)^n}{n!}\, \mathrm{d}t = M \frac{(x-a)^{n+1}}{(n+1)!}$$

On a utilisé $\left| \int_a^b f(t)\, \mathrm{d}t \right| \leqslant \int_a^b |f(t)|\, \mathrm{d}t$, valable pour $a \leqslant b$.

Si $a \geqslant x$, il faut remettre les bornes « dans le bon sens » :

$$\left| \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, \mathrm{d}t \right| = \left| \int_x^a \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, \mathrm{d}t \right| \leqslant M \int_x^a \frac{(t-x)^n}{n!}\, \mathrm{d}t = M \frac{(a-x)^{n+1}}{(n+1)!}$$

D'où la conclusion, dans les deux cas.

Applications

• L'inégalité de Taylor Lagrange permet de trouver les trois développements limités à retenir, voir § 1.1.5. Par exemple, pour l'exponentielle, cette formule fournit : ($x \in I$ segment contenant 0)

$$\left| \mathrm{e}^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \leqslant M \frac{|x|^{n+1}}{(n+1)!}$$

avec M majorant de e^x sur I. Ceci montre que $g(x) = \mathrm{e}^x - \sum_{k=0}^{n} \frac{x^k}{k!}$ est une fonction négligeable devant x^n quand x tend vers 0, car $\left| \frac{g(x)}{x^n} \right| \leqslant M \frac{|x|}{(n+1)!}$. On a donc bien, pour tout $n \in \mathbf{N}$:

$$\mathrm{e}^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \circ(x^n)$$

• Développement en série de e^x. Appliquons l'inégalité de Taylor-Lagrange avec $f(x) = e^x$, $a = 0$. Si $x \geqslant 0$, $|f^{(n+1)}(t)| = e^t$ est majoré par e^x sur $[0, x]$, donc

$$\left| e^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \leqslant e^x \frac{x^{n+1}}{(n+1)!}$$

Et il suffit de montrer que $\left(\frac{x^n}{n!}\right)$ converge vers 0 pour conclure. Or, pour $x \geqslant 0$ fixé, il existe $n_0 \in \mathbf{N}$ tel que $\frac{x}{n_0} \leqslant \frac{1}{2}$. Donc, pour tout $n > n_0$, on a $0 \leqslant \frac{x}{n} \leqslant \frac{1}{2}$, donc

$$0 \leqslant \frac{x^n}{n!} = \frac{x^{n_0}}{n_0!} \times \frac{x^{n-n_0}}{(n_0+1)\dots n} \leqslant \frac{x^{n_0}}{n_0!} \times \left(\frac{1}{2}\right)^{n-n_0} \xrightarrow[n\to+\infty]{} 0$$

D'où la conclusion dans ce cas là. Si $x < 0$, alors $|f^{(n+1)}(t)| = e^t$ est majoré par 1 sur $]-\infty, 0]$, et la suite est analogue. On a bien établi :

$$\forall x \in \mathbf{R},\ e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$$

• La formule de Taylor avec reste intégral appliquée à la fonction exponentielle permet d'établir :

$$\forall x \in [0, +\infty[\,, \forall n \in \mathbf{N}, \quad e^x \geqslant \sum_{k=0}^{n} \frac{x^k}{k!}\ ;$$

$$\forall x \in \,]-\infty, 0]\,, \forall p \in \mathbf{N}^*, \quad \sum_{k=0}^{2p-1} \frac{x^k}{k!} \leqslant e^x \leqslant \sum_{k=0}^{2p} \frac{x^k}{k!}.$$

Il suffit d'étudier le signe du reste intégral.

3. Intégrales généralisées

Si la fonction f est continue sur le segment $[a, b]$, on sait définir l'intégrale $\int_a^b f(t)\,dt$. Cette définition peut être dans certains cas étendue.

3.1 Intégrale d'une fonction continue par morceaux sur $[a, b]$

Définition. On dit que la fonction f est **continue par morceaux** sur l'intervalle I ssi f est continue sauf en un nombre fini de points, où f admet une limite finie à droite et à gauche.

> **Théorème.** Soit f une fonction continue sur l'intervalle $[a, b[$, et prolongeable par continuité en b. Alors la fonction $x \mapsto \int_a^x f(t)\, dt$ admet une limite finie en b, et on pose :
>
> $$\int_a^b f(t)\, dt = \lim_{x \to b,\ x<b} \int_a^x f(t)\, dt$$

On a un théorème analogue pour f continue sur $]a, b]$, prolongeable par continuité en a.

Ce théorème permet de définir l'intégrale $\int_a^b f(t)\, dt$, avec f continue par morceaux sur l'intervalle $[a, b]$. Supposons en effet, pour fixer les idées, que f soit continue sur $[a, b]$ sauf en c, où f admet une limite finie à gauche et à droite. Alors on définit l'intégrale $\int_a^b f(t)\, dt$ en utilisant la relation de Chasles pour poser :

$$\int_a^b f(t)\, dt = \int_a^c f(t)\, dt + \int_c^b f(t)\, dt$$

Chacune des intégrales du membre de droite existe, d'après le théorème précédent.

> Intégrale $\int_0^3 \mathrm{Ent}(x)\, dx$. La fonction partie entière Ent (§1.5.3) est continue par morceaux sur $[0, 3]$, donc l'intégrale existe, et
>
> $$\int_0^3 \mathrm{Ent}(x)\, dx = \int_0^1 0\, dx + \int_1^2 1\, dx + \int_2^3 2\, dx = 0 + 1 + 2 = 3$$

3.2 Intégrale d'une fonction continue sur $]a, b]$ ou sur $[a, b[$

L'existence de l'intégrale n'est alors plus assurée.

Définitions

Soit $a, b \in \mathbf{R}$, $a < b$, et f une fonction continue sur $]a, b]$. On pose

$$\int_a^b f(t)\, dt = \lim_{x \to a,\ x>a} \int_x^b f(t)\, dt \text{ sous réserve de limite finie,}$$

et on dit alors que **l'intégrale généralisée** (ou **impropre**) $\int_a^b f(t)\, dt$ est **convergente**. Dans le cas contraire, on dit qu'elle est **divergente**.

On définit de même, avec f continue sur $[a, b[$:

$$\int_a^b f(t)\, dt = \lim_{x \to b,\ x<b} \int_a^x f(t)\, dt \text{ sous réserve de limite finie}$$

Critères de convergence

- **Intégrales de Riemann**

$$\int_0^1 \frac{1}{t^\alpha} \mathrm{d}t \text{ est convergente ssi } \alpha < 1$$

Montrons par exemple que l'intégrale $\int_0^1 \frac{1}{\sqrt{t}}\, \mathrm{d}t$ est convergente. La fonction $t \mapsto \frac{1}{\sqrt{t}}$ est continue sur $]0, 1]$, et, avec $0 < x < 1$:

$$\int_x^1 \frac{1}{\sqrt{t}}\, \mathrm{d}t = \left[2\sqrt{t}\right]_x^1 = 2 - 2\sqrt{x} \xrightarrow[x \to 0]{} 2$$

L'intégrale $\int_0^1 \frac{1}{\sqrt{t}}\, \mathrm{d}t$ est donc convergente, et $\int_0^1 \frac{1}{\sqrt{t}}\, \mathrm{d}t = 2$.

- **Cas des fonctions positives**

$$\left.\begin{array}{c} f \geqslant 0, g \geqslant 0 \text{ sur } [a, +\infty[\\ f \underset{+\infty}{\sim} g \end{array}\right\} \Rightarrow \int_a^{+\infty} f(t)\, \mathrm{d}t \text{ et } \int_a^{+\infty} g(t)\, \mathrm{d}t \text{ sont}$$

de même nature, convergente ou divergente

Intégrale $\int_0^1 \frac{t+1}{\sqrt{t}} \mathrm{d}t$. La fonction $t \mapsto \frac{t+1}{\sqrt{t}}$ est positive et continue sur $]0, 1]$; En 0, $\frac{t+1}{\sqrt{t}} \sim \frac{1}{\sqrt{t}}$, et l'intégrale $\int_0^1 \frac{1}{\sqrt{t}}\, \mathrm{d}t$ est une intégrale de Riemann convergente. Donc l'intégrale $\int_0^1 \frac{t+1}{\sqrt{t}} \mathrm{d}t$ est convergente.

3.3 Intégrale sur un intervalle $[a, +\infty[$, $]-\infty, b]$

Définitions

Soit $a \in \mathbf{R}$, et f une fonction continue sur $[a, +\infty[$. On pose

$$\int_a^{+\infty} f(t)\, \mathrm{d}t = \lim_{X \to +\infty} \int_a^X f(t)\, \mathrm{d}t \text{ sous réserve de limite finie,}$$

et on dit alors que **l'intégrale généralisée** (ou **impropre**) $\int_a^{+\infty} f(t)\, \mathrm{d}t$ est **convergente**. Dans le cas contraire, on dit qu'elle est **divergente**.
On définit de même, avec f continue sur $]-\infty, b]$:

$$\int_{-\infty}^b f(t)\, \mathrm{d}t = \lim_{X \to -\infty} \int_{-\infty}^b f(t)\, \mathrm{d}t \text{ sous réserve de limite finie}$$

Intégrale $\int_0^{+\infty} \frac{x}{(x^2+1)^2}\,\mathrm{d}x$. La fonction $x \mapsto \frac{x}{(x^2+1)^2}$ est continue sur $[0, +\infty[$, et

$$\int_0^X \frac{x}{\left(x^2+1\right)^2}\,\mathrm{d}x = \left[-\frac{1}{2} \times \frac{1}{x^2+1}\right]_0^X = -\frac{1}{2\left(X^2+1\right)} + \frac{1}{2} \xrightarrow[X\to+\infty]{} \frac{1}{2}$$

L'intégrale est donc convergente, et

$$\int_0^{+\infty} \frac{x}{\left(x^2+1\right)^2}\,\mathrm{d}x = \frac{1}{2}$$

Critères de convergence

• **Intégrales de Riemann**

$$\text{L'intégrale } \int_1^{+\infty} \frac{1}{t^\alpha}\,\mathrm{d}t \text{ converge ssi } \alpha > 1$$

En effet, la fonction $t \mapsto \frac{1}{t^\alpha}$ est continue sur $[1, +\infty[$, et :

Si $\alpha \neq 1$, $\int_1^X \frac{1}{t^\alpha}\,\mathrm{d}t = \int_1^X t^{-\alpha}\,\mathrm{d}t = \left[\frac{t^{-\alpha+1}}{-\alpha+1}\right]_1^X = \frac{X^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha}$ tend vers une limite finie en $+\infty$ ssi $\alpha > 1$. Avec

$$\alpha = 1, \int_1^X \frac{1}{t}\,\mathrm{d}t = \ln X \xrightarrow[X\to+\infty]{+} \infty.$$

Attention au comportement différent des intégrales de Riemann suivant l'intervalle :

$\int_0^1 \frac{1}{t^\alpha}\,\mathrm{d}t$ est convergente ssi $\alpha < 1$;

$\int_1^{+\infty} \frac{1}{t^\alpha}\,\mathrm{d}t$ est convergente ssi $\alpha > 1$.

• **Cas des fonctions positives**

$$\left.\begin{array}{l} 0 \leqslant f \leqslant g \text{ sur } [a, +\infty[\\ \int_a^{+\infty} g(t)\,\mathrm{d}t \text{ convergente} \end{array}\right\} \Rightarrow \int_a^{+\infty} f(t)\,\mathrm{d}t \text{ convergente}$$

$$\left.\begin{array}{l} 0 \leqslant f \leqslant g \text{ sur } [a, +\infty[\\ \int_a^{+\infty} f(t)\,\mathrm{d}t \text{ divergente} \end{array}\right\} \Rightarrow \int_a^{+\infty} g(t)\,\mathrm{d}t \text{ divergente}$$

$$\left.\begin{array}{c} f \geqslant 0, g \geqslant 0 \text{ sur } [a, +\infty[\\ f \underset{+\infty}{\sim} g \end{array}\right\} \Rightarrow \int_a^{+\infty} f(t)\,\mathrm{d}t \text{ et } \int_a^{+\infty} g(t)\,\mathrm{d}t \text{ de}$$

même nature, convergente ou divergente

$$\left.\begin{array}{c} f \geqslant 0, g \geqslant 0 \text{ sur } [a, +\infty[\\ f \underset{+\infty}{=} \circ(g) \\ \int_a^{+\infty} g(t)\,\mathrm{d}t \text{ convergente} \end{array}\right\} \Rightarrow \int_a^{+\infty} f(t)\,\mathrm{d}t \text{ convergente}$$

Pour démontrer le premier critère, on considère $\varphi(x) = \int_a^x f(t)\,\mathrm{d}t$. Sur $[a, +\infty[$, φ est croissante, car $\varphi'(x) = f(x) \geqslant 0$, et majorée, car $\varphi(x) \leqslant \int_a^x g(t)\,\mathrm{d}t \leqslant \int_a^{+\infty} g(t)\,\mathrm{d}t = M \in \mathbf{R}$. φ admet donc une limite finie en $+\infty$, ce qu'il fallait démontrer.

On a les critères analogues pour les intégrales $\int_{-\infty}^b f(t)\,\mathrm{d}t$.

• **Fonctions de signe quelconque**

Définition. On dit que l'intégrale $\int_a^{+\infty} f(t)\,\mathrm{d}t$ est **absolument convergente** ssi l'intégrale $\int_a^{+\infty} |f(t)|\,\mathrm{d}t$ est convergente.

Théorème. Si l'intégrale $\int_a^{+\infty} f(t)\,\mathrm{d}t$ est absolument convergente, alors elle est convergente.

On a le théorème analogue pour l'intégrale $\int_{-\infty}^b f(t)\,\mathrm{d}t$.

Remarque. On a alors $\left|\int_a^{+\infty} f(t)\,\mathrm{d}t\right| \leqslant \int_a^{+\infty} |f(t)|\,\mathrm{d}t$, et de même pour $\int_{-\infty}^b f(t)\,\mathrm{d}t$.

3.4 Cas général — Calcul d'intégrales généralisées

Cas général

Dans le cas général, avec $-\infty \leqslant a < c < b \leqslant +\infty$, on dira que l'intégrale $\int_a^b f(t)\,\mathrm{d}t$ est convergente ssi chacune des intégrales $\int_a^c f(t)\,\mathrm{d}t$, $\int_c^b f(t)\,\mathrm{d}t$ l'est, et on a alors (relation de Chasles) :

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$$

ce qui permet de proche en proche de se ramener à un des cas étudiés précédemment, étant entendu qu'une intégrale définie est convergente.

Intégrale $\int_0^{+\infty} \frac{1}{t^2+1}\,\mathrm{d}t$. La fonction $t \mapsto \frac{1}{1+t^2}$ est continue et positive sur $[0, +\infty[$; $\frac{1}{t^2+1} \underset{+\infty}{\sim} \frac{1}{t^2}$ (ou bien : pour tout $t \geqslant 1$, $\frac{1}{t^2+1} \leqslant \frac{1}{t^2}$) ;

L'intégrale $\int_1^{+\infty} \frac{1}{t^2}\, \mathrm{d}t$ est une intégrale de Riemann convergente.
Donc l'intégrale $\int_1^{+\infty} \frac{1}{t^2+1}\, \mathrm{d}t$ est convergente, et d'après la relation de Chasles, l'intégrale $\int_0^{+\infty} \frac{1}{t^2+1}\, \mathrm{d}t = \int_0^1 \frac{1}{t^2+1}\, \mathrm{d}t + \int_1^{+\infty} \frac{1}{t^2+1}\, \mathrm{d}t$ est aussi convergente ($\int_0^1 \frac{1}{t^2+1}\, \mathrm{d}t$ est une intégrale définie.)

Utilisation de la parité

Soit f une fonction définie sur $\mathbf{R}$ et telle que $\int_0^{+\infty} f(t)\, \mathrm{d}t$ converge.

- Si f est paire, alors $\int_{-\infty}^{+\infty} f(t)\, \mathrm{d}t$ converge, et

$$\int_{-\infty}^{+\infty} f(t)\, \mathrm{d}t = 2\int_0^{+\infty} f(t)\, \mathrm{d}t$$

- Si f est impaire, alors $\int_{-\infty}^{+\infty} f(t)\, \mathrm{d}t$ converge, et

$$\int_{-\infty}^{+\infty} f(t)\, \mathrm{d}t = 0$$

Intégrale $\int_{-\infty}^{+\infty} \exp\left(-t^2/2\right)\, \mathrm{d}t$. $f : t \mapsto \exp\left(-t^2/2\right)$ est continue et positive ou nulle sur $[0, +\infty[$, négligeable devant $\frac{1}{t^2}$ quand t tend vers $+\infty$. $\int_1^{+\infty} \frac{1}{t^2}\, \mathrm{d}t$ est convergente, donc $\int_0^{+\infty} \exp\left(-t^2/2\right)\, \mathrm{d}t$ est convergente. f est paire, donc l'intégrale $\int_{-\infty}^{+\infty} \exp\left(-t^2\right)\, \mathrm{d}t$ est convergente, et $\int_{-\infty}^{+\infty} \exp\left(-t^2\right)\, \mathrm{d}t = 2\int_0^{+\infty} \exp\left(-t^2\right)\, \mathrm{d}t$

Calculs d'intégrales généralisées

- **Utilisation de la définition.** Soit $I = \int_1^{+\infty} \frac{\mathrm{e}^{\frac{1}{x}}}{x^2}\, \mathrm{d}x$

La fonction $f : x \mapsto \frac{1}{x^2}\exp\left(\frac{1}{x}\right)$ est continue sur $[1, +\infty[$, donc

$$I = \lim_{X\to+\infty} \int_1^X \frac{\mathrm{e}^{\frac{1}{x}}}{x^2}\, \mathrm{d}x \text{ sous réserve de limite finie.}$$

Or $\displaystyle\int_1^X \frac{\mathrm{e}^{\frac{1}{x}}}{x^2}\, \mathrm{d}x = \left[-\mathrm{e}^{\frac{1}{x}}\right]_1^X = -\mathrm{e}^{\frac{1}{X}} + \mathrm{e} \xrightarrow[X\to+\infty]{} \mathrm{e} - 1$

Donc l'intégrale I est convergente, et vaut $\mathrm{e} - 1$.

- **Utilisation de la définition et d'une relation de récurrence.** Soit $I_n = \int_0^{+\infty} t^n \mathrm{e}^{-t}\, \mathrm{d}t$, $n \in \mathbf{N}$. On peut prouver que l'intégrale I_n est

convergente : la fonction $t \mapsto t^n e^{-t}$ est positive et continue sur $\mathbf{R}^+$, négligeable en $+\infty$ devant $\frac{1}{t^2}$, et l'intégrale $\int_1^{+\infty} \frac{1}{t^2}\,dt$ est convergente. Pour le calcul, on utilise la définition pour déterminer I_0 :

$$\int_0^X e^{-t}\,dt = \left[-e^{-t}\right]_0^X = -e^{-X} + 1 \underset{X \to +\infty}{\longrightarrow} I_0 = 1,$$

puis une intégration par parties portant sur l'intégrale $\int_0^X t^n e^{-t}\,dt$ avec $X > 0$; $u = t^n$, $v' = e^{-t}$, $u' = nt^{n-1}$, $v = -e^{-t}$; u, v de classe C^1. En passant à la limite $X \to +\infty$ dans la relation trouvée, on obtient $I_n = nI_{n-1}$, $n \in \mathbf{N}^*$, et finalement, par récurrence : $\forall n \in \mathbf{N}, I_n = n!$.

• **Utilisation d'un changement de variable affine.** On admet le résultat :

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt = \sqrt{2\pi} \qquad \text{(à retenir, voir § 9.2.3)}$$

On considère alors, avec $m \in \mathbf{R}, \sigma > 0$:

$$I_{m,\sigma} = \int_{-\infty}^{+\infty} e^{-\frac{(t-m)^2}{2\sigma^2}}\,dt$$

Soit $y = \frac{t-m}{\sigma}$. On a $dt = \sigma\,dy$. σ est positif, donc

$$I_{m,\sigma} = \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}}\sigma\,dt$$

L'intégrale $I_{m,\sigma}$ est donc convergente, et vaut $\sigma\sqrt{2\pi}$.

En dehors de ce cas (changement de variable affine), tous les techniques de calcul d'intégrale seront pratiquées sur des intégrales définies.

4. Séries et intégrales

Comparaison d'une série et d'une intégrale généralisée

Théorème. Si f est une fonction continue, positive et décroissante sur $[1, +\infty[$, alors la série $\sum_{n\in\mathbf{N}^*} f(n)$ et l'intégrale $\int_1^{+\infty} f(t)\,dt$ sont de même nature, divergente ou convergente.

Éléments pour la démonstration : $\forall k \in \mathbf{N}^*, f(k+1) \leqslant f(x) \leqslant f(k)$; en intégrant sur $[k, k+1]$ cette double inégalité, on obtient :

$$(1) \quad f(k+1) \leqslant \int_k^{k+1} f(x)\,\mathrm{d}x \leqslant f(k)$$

On additionne de $k = 1$ à $k = n - 1$ les inégalités de gauche de (1), puis on ajoute $f(1)$ aux deux membres de l'inégalité obtenue. On additionne de $k = 1$ à $k = n$ les inégalités de droite de (1). En remettant dans le bon ordre les résultats obtenus, on obtient

$$\int_1^{n+1} f(x)\,\mathrm{d}x \leqslant \sum_{k=1}^{n} f(k) \leqslant f(1) + \int_1^{n} f(x)\,\mathrm{d}x$$

ce qui permet de conclure par passage à la limite.

L'intégrale de Riemann $\int_1^{+\infty} \frac{1}{x^\alpha}\,\mathrm{d}x$ est convergente ssi $\alpha > 1$, donc la série de Riemann $\sum_{n \in \mathbf{N}^*} \frac{1}{n^\alpha}$ est convergente sssi $\alpha > 1$.

Un autre exemple

Soit t, x tels que $0 \leqslant t \leqslant x \leqslant 1$. L'identité géométrique fournit :

$$\frac{1}{1+t} - \sum_{k=0}^{n} (-1)^k t^k = \frac{(-t)^{n+1}}{1+t}$$

En intégrant sur l'intervalle $[0, x]$, et en utilisant les propriétés de l'intégrale (linéarité, calculs d'intégrales, intégrale et valeur absolue), on obtient :

$$\int_0^x \frac{1}{1+t}\,\mathrm{d}t - \sum_{k=0}^{n} (-1)^k \int_0^x t^k\,\mathrm{d}t = \int_0^x \frac{(-t)^{n+1}}{1+t}\,\mathrm{d}t$$

$$\left| \ln(1+x) - \sum_{k=0}^{n} (-1)^k \frac{x^{k+1}}{k+1} \right| \leqslant \int_0^x \frac{t^{n+1}}{1+t}\,\mathrm{d}t \leqslant \int_0^x t^{n+1}\,\mathrm{d}t = \frac{x^{n+2}}{n+2}$$

Avec $x = 1$, par passage à la limite, on obtient :

$$\ln 2 = \sum_{k=0}^{+\infty} \frac{(-1)^k}{k+1}$$

Partie 2

Algèbre linéaire



Systèmes linéaires Calcul matriciel

4

1. Systèmes linéaires

1.1 Généralités

Définitions. Un **système d'équations linéaires** est un système d'équations du type

$$(1)\begin{cases} a_{1,1}x + a_{1,2}y + a_{1,3}z = b_1 & L_1 \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = b_2 & L_2 \\ a_{3,1}x + a_{3,2}y + a_{3,3}z = b_3 & L_3 \end{cases}$$

Ici il s'agit d'un système de 3 équations L_1, L_2, L_3, à 3 inconnues, les inconnues x, y, z appartenant à $\mathbf{R}$.

Résoudre le système (1), c'est déterminer l'ensemble de ses solutions S, c'est-à-dire l'ensemble des valeurs de (x, y, z) tels que les trois égalités L_1, L_2, L_3 soient simultanément vraies. S est une partie de $\mathbf{R}^3$, éventuellement vide.

Deux systèmes sont **équivalents** ssi ils ont même ensemble de solutions.

Remarque. On peut toujours se ramener au cas où le nombre d'équations est égal au nombre d'inconnues. Par exemple :

$$\begin{cases} a_{1,1}x + a_{1,2}y = b_1 & L_1 \\ a_{2,1}x + a_{2,2}y = b_2 & L_2 \\ a_{3,1}x + a_{3,2}y = b_3 & L_3 \end{cases}$$

On résout le système formé par L_1, L_2, puis on reporte dans L_3.

$$\begin{cases} a_{1,1}x + a_{1,2}y + a_{1,3}z = b_1 \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = b_2 \end{cases}$$

On résout

$$\begin{cases} a_{1,1}x + a_{1,2}y = b_1 - a_{1,3}z \\ a_{2,1}x + a_{2,2}y = b_2 - a_{2,3}z \end{cases}$$

1.2 Méthode du pivot

Pour fixer les idées, on s'intéresse au système (1) – trois équations, trois inconnues. Les principes de la méthode du pivot sont les suivants :

- Les systèmes les plus faciles à résoudre sont les **systèmes triangulaires**, ou **diagonaux**, c'est-à-dire de la forme

$$(2)\begin{cases}\alpha_{1,1}x+\alpha_{1,2}y+\alpha_{1,3}z=\beta_1\\ \alpha_{2,2}y+\alpha_{2,3}z=\beta_2\\ \alpha_{3,3}z=\beta_3\end{cases}\quad (2')\begin{cases}\alpha_{1,1}x=\beta_1\\ \alpha_{2,2}y=\beta_2\\ \alpha_{3,3}z=\beta_3\end{cases}$$

- Il est toujours possible d'obtenir un système triangulaire (2) équivalent au système de départ (1) en effectuant les **opérations élémentaires sur les lignes** :
  $L_i \leftrightarrow L_j$: échange de la ligne L_i et de la ligne L_j.
  $L_i \leftarrow \alpha L_i + \beta L_j$: on remplace la ligne L_i par $\alpha L_i + \beta L_j$.
  β peut être nul, mais **ATTENTION !** α **doit être non nul !**
- Les coefficients diagonaux $\alpha_{1,1}, \alpha_{2,2}, \alpha_{3,3}$ du système triangulaire obtenu (2) sont appelés **pivots** du système (1). On résout ce système de proche en proche.

Exemple :

$$\begin{cases}\mathbf{2}x_1-4x_2+x_3=5\\ x_1+x_2+x_3=2 & L_2\leftarrow \mathbf{2}L_2-L_1\\ -x_1+4x_2-x_3=-2 & L_3\leftarrow \mathbf{2}L_3+L_1\end{cases}$$

2 est le premier **pivot**. On élimine l'inconnue du pivot des autres équations en effectuant $L_i \leftarrow \mathbf{2}L_i + \beta L_1$.

$$\begin{cases}2x_1-4x_2+x_3=5\\ \mathbf{6}x_2+x_3=-1\\ 4x_2-x_3=1 & L_3\leftarrow 3L_3-2L_2\end{cases}$$

6 deuxième pivot. $L_3 \leftarrow 3L_3 - 2L_2$ préférable à $L_3 \leftarrow 6l_3 - 4L_2$.

$$\begin{cases}2x_1-4x_2+x_3=5\\ 6x_2+x_3=-1\\ -\mathbf{5}x_3=5\end{cases}$$

On a obtenu un système triangulaire. $-\mathbf{5}$ est le dernier pivot. On trouve $x_3=-1$, puis $x_2=0$, puis $x_1=3$.

L'ensemble des solutions est $\{(3,0,-1)\}$.

1.3 Systèmes de Cramer, systèmes homogènes

Définitions. Un système linéaire est dit **de Cramer** ssi il admet une solution unique.

Un système linéaire est dit **homogène** ssi les seconds membres des équations qui le composent sont tous égaux à 0 :

$$(3)\begin{cases} a_{1,1}x + a_{1,2}y + a_{1,3}z = 0 \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = 0 \\ a_{3,1}x + a_{3,2}y + a_{3,3}z = 0 \end{cases}$$

Propriétés.

- Un système triangulaire (2) admet une solution unique ssi ses coefficients diagonaux sont nuls. La méthode du pivot montre alors :

Un système est de Cramer ssi tous ses pivots sont non nuls.

On remarque que cette propriété ne dépend pas des seconds membres des équations du système.

- Propriétés des systèmes homogènes :
  – Le système (3) admet toujours la solution $(0, 0, 0)$.
  – Soient $u = (x, y, z)$ et $u' = \left(x', y', z'\right)$ solutions du système (3), et a, b deux nombres réels.
  Alors la **somme** :
  $$u + u' = \left(x + x', y + y', z + z'\right),$$
  Le **produit de** u **par le nombre réel** a :
  $$au = (ax, ay, az),$$
  Et plus généralement les **combinaisons linéaires** de u, u' :
  $$au + bu' = \left(ax + bx', ay + by', az + bz'\right),$$
  sont solutions du système (3).
  Réciproquement, quand un système homogène admet des solutions non nulles, on exprimera celles-ci comme des combinaisons linéaires de solutions fixées. Voir l'application fondamentale.

Une application fondamentale.

- La recherche de valeurs propres et vecteurs propres d'une matrice ou d'un endomorphisme (voir chap 5 et 6) conduit à la résolution de systèmes du type

$$\begin{cases} a_{1,1}x + a_{1,2}y + a_{1,3}z = \lambda x \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = \lambda y \\ a_{3,1}x + a_{3,2}y + a_{3,3}z = \lambda z \end{cases} \Leftrightarrow \begin{cases} \left(a_{1,1} - \lambda\right) x + a_{1,2}y + a_{1,3}z = 0 \\ a_{2,1}x + \left(a_{2,2} - \lambda\right) y + a_{2,3}z = 0 \\ a_{3,1}x + a_{3,2}y + \left(a_{3,3} - \lambda\right) z = 0 \end{cases}$$

On doit déterminer l'ensemble des solutions en discutant suivant la valeur de $\lambda \in \mathbf{R}$.

Étude détaillée d'un exemple.

$$(4)\begin{cases} -\lambda x + \ \ y + \ \ z = 0 \\ \ \ x - \lambda y + \ \ z = 0 \qquad L_1 \leftrightarrow L_2 \\ \ \ x + \ \ y - \lambda z = 0 \end{cases}$$

On ne peut pas prendre $-\lambda$ comme premier pivot, car on serait amené aux opérations élémentaires $L_2 \leftarrow \lambda L_2 + L_1$; $L_3 \leftarrow \lambda L_3 + L_1$ et l'équivalence est perdue si $\lambda = 0$. Une discussion prématurée est à proscrire. L'opération $L_1 \leftrightarrow L_2$ permet de prendre **1** pour pivot :

$$\begin{cases} \mathbf{1}x - \lambda y + \ \ z = 0 \\ -\lambda x + \ \ y + \ \ z = 0 \qquad L_2 \leftarrow L_2 + \lambda L_1 \\ \ \ x + \ \ y - \lambda z = 0 \qquad L_3 \leftarrow L_3 - L_1 \end{cases}$$

$$(5)\begin{cases} x - \lambda y + z = 0 \\ \left(1 - \lambda^2\right) y + (1 + \lambda) z = 0 \\ (1 + \lambda) y - (\lambda + 1) z = 0 \end{cases}$$

De même, on ne peut pas prendre $1 - \lambda^2$ comme pivot. On a alors le choix entre **trois possibilités** :

Première possibilité : on effectue sur le système (5) l'opération élémentaire $L_2 \leftrightarrow L_3$:

$$\begin{cases} x - \lambda y + z = 0 \\ (1 + \lambda) y - (\lambda + 1) z = 0 \\ \left(1 - \lambda^2\right) y + (1 + \lambda) z = 0 \end{cases}$$

Puis on effectue $L_3 \leftarrow L_3 - (1 - \lambda) L_2$:

$$\begin{cases} x - \lambda y + z = 0 \\ (1 + \lambda) y - (\lambda + 1) z = 0 \\ (1 + \lambda)(2 - \lambda) z = 0 \end{cases}$$

On a obtenu un système triangulaire, et on peut commencer la discussion. Soit S_λ l'ensemble des solutions.

– Si $\lambda \notin \{-1, 2\}$, le système est de Cramer, donc $S_\lambda = \{(0, 0, 0)\}$.

– Si $\lambda = -1$: (3) $\Leftrightarrow x + y + z = 0 \Leftrightarrow x = -y - z$.

Donc

$$\begin{aligned} (x, y, z) \in S_{-1} \Leftrightarrow (x, y, z) &= (-y - z, y, z) \\ &= y(-1, 1, 0) + z(-1, 0, 1). \end{aligned}$$

Donc $\quad S_{-1} = \{y(-1, 1, 0) + z(-1, 0, 1) \ ; y, z \in \mathbf{R}\}$.

– Si $\lambda = 2$: (4) $\Leftrightarrow \begin{cases} x - 2y + \quad z = 0 \\ \quad 3y - 3z = 0 \end{cases} \Leftrightarrow \begin{cases} x - 2y = -z \\ \quad\quad y = \quad z \end{cases} \begin{cases} x = z \\ y = z \end{cases}$

Donc (x, y, z) solution $\Leftrightarrow (x, y, z) = (z, z, z) = z\,(1, 1, 1)$.

Dans les cas où S_λ n'est pas réduit à la solution nulle, on a bien obtenu les solutions comme combinaisons linéaires de solutions fixes.

Deuxième possibilité : on effectue sur le système (5) l'opération $L_3 \leftarrow L_3 + L_2$:

$$\begin{cases} x - \lambda y + z = 0 \\ \left(1 - \lambda^2\right) y + (1 + \lambda)\, z = 0 \\ \left(2 + \lambda - \lambda^2\right) y = 0 \end{cases}$$

On obtient un système qui n'est pas triangulaire, mais **échelonné**, avec une équation à une, à deux, à trois inconnues. On peut commencer la discussion : si $2 + \lambda - \lambda^2 \neq 0 \Leftrightarrow \lambda \notin \{-1, 2\}$, alors $y = 0$, puis $z = 0$, puis $x = 0$. Si $\lambda = -1$, le système se réduit à l'équation $x + y + z = 0$, etc.

Troisième possibilité : dans le système (5), on permute la place des inconnues y, z :

$$\begin{cases} x + z - \lambda y = 0 \\ -(\lambda + 1)\, z + (1 + \lambda)\, y = 0 \\ (1 + \lambda)\, z + \left(1 - \lambda^2\right) y = 0 \end{cases}$$

puis on effectue l'opération élémentaire $L_3 \leftarrow L_3 + L_1$ pour obtenir un système triangulaire, que l'on résout comme précédemment.

- Afin d'alléger l'écriture et les calculs, on peut adopter une notation simplifiée pour les systèmes homogènes : on ne garde que les coefficients des inconnues que l'on écrit dans une **matrice** (voir § suivant), le reste (emplacement des inconnues, seconds membres) étant sans changement d'étape en étape. Pour le système étudié en exemple on écrira donc :

$$\begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} \qquad L_1 \leftrightarrow L_2$$

$$\begin{pmatrix} 1 & -\lambda & 1 \\ -\lambda & 1 & 1 \\ 1 & 1 & -\lambda \end{pmatrix} \qquad \begin{matrix} L_2 \leftarrow L_2 + \lambda L_1 \\ L_3 \leftarrow L_3 - L_1 \end{matrix}$$

$$\begin{pmatrix} 1 & -\lambda & 1 \\ 0 & 1 - \lambda^2 & 1 + \lambda \\ 0 & 1 + \lambda & -(\lambda + 1) \end{pmatrix}$$

etc. Les opérations indiquées se font sur les **lignes** des matrices.

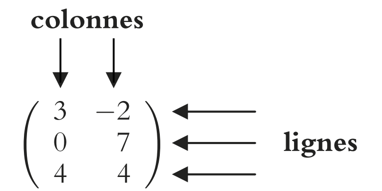
Une fois le système triangulaire ou échelonné obtenu, il est préférable de retourner à la notation traditionnelle.

On se gardera d'utiliser cette notation si on a choisi de permuter la place des inconnues. Toute manipulation sur les **colonnes** d'une matrice est exclue.

2. Calcul matriciel

2.1 Définitions

Définitions générales. Une **matrice** (on précise quelquefois **matrice réelle**) est un tableau de nombres réels. Exemple d'une matrice à 3 lignes et 2 colonnes :



Les nombres du tableau sont appelés les **coefficients**, ou **termes** de la matrice. Le terme en deuxième ligne et première colonne est égal à 0.

Pour n, p appartenant à $\mathbf{N}^*$, on note $\mathcal{M}_{n,p}(\mathbf{R})$ **l'ensemble des matrices à** n **lignes et** p **colonnes**. La matrice ci-dessus appartient à $\mathcal{M}_{3,2}(\mathbf{R})$.

La **matrice nulle** $O_{n,p}$ est la matrice de $\mathcal{M}_{n,p}(\mathbf{R})$ dont tous les termes sont nuls.

Une matrice appartenant à $\mathcal{M}_{1,p}(\mathbf{R})$ s'appelle **matrice-ligne**.

Une matrice appartenant à $\mathcal{M}_{n,1}(\mathbf{R})$ s'appelle **matrice-colonne**.

L'ensemble $\mathcal{M}_{n,n}(\mathbf{R})$ est simplement noté $\mathcal{M}_n(\mathbf{R})$. Ses éléments sont appelés **matrices carrées d'ordre** n.

$$(1\ \ 2\ \ 3) \in \mathcal{M}_{1,3}(\mathbf{R})\ ;\ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbf{R})\ ;\ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathcal{M}_3(\mathbf{R}).$$

Matrices carrées. Soit $M = \left(a_{i,j}\right)_{1\leqslant i,j\leqslant n}$ une matrice carrée d'ordre n : $a_{i,j}$ désigne le coefficient en i-ème ligne et j-ème colonne.

La suite des nombres $\left(a_{i,i}\right)_{1\leqslant i\leqslant n}$ est la **diagonale** de la matrice M :

$$\begin{pmatrix} \mathbf{a_{1,1}} & a_{1,2} & a_{1,3} \\ a_{2,1} & \mathbf{a_{2,2}} & a_{2,3} \\ a_{3,1} & a_{3,2} & \mathbf{a_{3,3}} \end{pmatrix}$$

diagonale

Une **matrice diagonale** D est une matrice carrée dont tous les termes en dehors de la diagonale sont nuls.

$$D = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Une **matrice triangulaire supérieure** est une matrice carrée T_S dont tous les termes sous la diagonale sont nuls.

$$T_S = \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix}$$

Une **matrice triangulaire inférieure** est une matrice carrée T_I dont tous les termes au dessus de la diagonale sont nuls.

$$T_I = \begin{pmatrix} \times & 0 & 0 \\ \times & \times & 0 \\ \times & \times & \times \end{pmatrix}$$

Une **matrice triangulaire** est une matrice triangulaire supérieure ou inférieure.

Une **matrice symétrique** est une matrice carrée $\left(a_{i,j}\right)_{1 \leqslant i,j \leqslant n}$ telle que, pour tout i, j appartenant à $\{1, \ldots, n\}$: $a_{j,i} = a_{i,j}$.

$S = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$ est une matrice symétrique.

La symétrie doit s'entendre « par rapport à la diagonale ».

La **transposée** de la matrice $M = \left(a_{i,j}\right)_{1 \leqslant i,j \leqslant n}$ est la matrice ${}^tM$, définie par

$$ {}^tM = \left(b_{i,j}\right)_{1 \leqslant i,j \leqslant n} \ ; \forall i, j \in \{1, \ldots, n\}, b_{i,j} = a_{j,i}$$

Avec $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, ${}^tM = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$.

Une matrice est symétrique ssi elle est égale à sa transposée.

La **matrice unité** de $\mathcal{M}_n(\mathbf{R})$ est la matrice I_n carrée d'ordre n, diagonale, et dont les termes de la diagonale sont tous égaux à 1 :

$$\mathrm{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2.2 Opérations sur les matrices

Définitions. Soit $A = \left(a_{i,j}\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$, $B = \left(b_{i,j}\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ deux matrices appartenant à $\mathcal{M}_{n,p}(\mathbf{R})$, et x un nombre réel.

La **somme** des matrices A, B est définie par

$$A + B = \left(a_{i,j} + b_{i,j}\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$$

Le **produit** de la matrice A **par le réel** x est défini par

$$xA = \left(xa_{i,j}\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$$

Quand le nombre de colonnes de A est égal au nombre de lignes de B, on définit le **produit des deux matrices** A, B (dans cet ordre) comme étant la matrice AB obtenue en effectuant les produits **ligne par colonne** : le produit de la i-ème ligne de A par la j-ème colonne de B donne le terme situé en$i$-ème et j-ème colonne de AB.

Exemple :

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} u & x \\ v & y \\ w & z \end{pmatrix} = \begin{pmatrix} au + bv + cw & ax + by + cz \\ du + ev + fw & dx + ey + fz \end{pmatrix}$$

Somme, produit par un réel : propriétés. Pour A, B, C appartenant à $\mathcal{M}_{n,p}(\mathbf{R})$, x, y appartenant à $\mathbf{R}$:

$$A + B \in \mathcal{M}_{n,p}(\mathbf{R}) \ ; \ xA \in \mathcal{M}_{n,p}(\mathbf{R})$$

$$A + (B + C) = (A + B) + C$$

$$A + B = B + A$$

$$A + \mathrm{O}_{n,p} = \mathrm{O}_{n,p} + A = A$$

$$A + (-A) = (-A) + A = \mathrm{O}_{n,p} \text{ en notant } -A \text{ la matrice } (-1)\,A$$

$$(x + y)\,A = xA + yA$$

$$x\,(A + B) = xA + xB$$

$$x\,(yA) = (xy)\,A$$

$$1A = A$$

Toutes ces propriétés sont des conséquences directes des propriétés connues de l'addition et de la multiplication des nombres réels. On les résume en disant que $\mathcal{M}_{n,p}(\mathbf{R})$ est un **espace vectoriel** sur $\mathbf{R}$.

De même que pour le nombres réels, on définit la différence de deux matrices par $A - B = A + (-B)$.

Produit de deux matrices : propriétés. Avec A, B, C matrices et x nombre réel, et chaque fois que toutes les opérations ont un sens :

$$A(BC) = (AB)C$$

$$A(B + C) = AB + AC$$

$$(A + B)C = AC + BC$$

$$x(AB) = (xA)B = A(xB)$$

On peut traduire cette dernière propriété en disant : pour multiplier un produit de matrices par un nombre réel, on multiplie une des matrices par ce nombre.

Attention, le produit BA peut ne pas exister alors que le produit AB existe. Ou bien ils peuvent être dans des ensembles $\mathcal{M}_{n,p}(\mathbf{R})$ différents. On bien, étant dans le même ensemble, ils peuvent ne pas être égaux.

Deux cas particuliers importants :

Produit de deux matrices carrées

Soit A, B, C trois matrices de $\mathcal{M}_n(\mathbf{R})$, x un nombre réel. Alors :

- $AB \in \mathcal{M}_n(\mathbf{R})$; $BA \in \mathcal{M}_n(\mathbf{R})$
- $A(BC) = (AB)C$;
- $A(B + C) = AB + AC$; $(A + B)C = AC + BC$;
- $x(AB) = (xA)B = A(xB)$;
- $I_nA = AI_n = A$;
- en général, $AB \neq BA$;
- on ne peut pas « simplifier » :

$$AB = AC \text{ n'implique pas } B = C, \text{ même si } A \neq \mathrm{O_n}.$$

Produit d'une matrice carrée et d'une matrice-colonne

Soit A, B deux matrices de $\mathcal{M}_n(\mathbf{R})$; X, Y deux matrices de $\mathcal{M}_{n,1}(\mathbf{R})$; x un nombre réel. Alors

- $AX \in \mathcal{M}_{n,1}(\mathbf{R})$;
- $I_nX = X$;
- $A(X + Y) = AX + AY$; $(A + B)X = AX + BX$;
- $A(xX) = (xA)X = x(AX)$.

2.3 Matrices carrées inversibles

Définition. Soit $A \in \mathcal{M}_n(\mathbf{R})$. On dit que A est **inversible** ssi il existe une matrice notée A^{-1} appartenant à $\mathcal{M}_n(\mathbf{R})$ telle que

$$AA^{-1} = A^{-1}A = I_n.$$

La matrice A^{-1} est la **matrice inverse** de la matrice A.
Si A, B sont inversibles, alors AB et A^{-1} sont inversibles, et

$$(AB)^{-1} = B^{-1}A^{-1} \ ; \ \left(A^{-1}\right)^{-1} = A.$$

Théorème. Soient A, B deux matrices de $\mathcal{M}_n(\mathbf{R})$ telles que $AB = I_n$. Alors A est inversible et $A^{-1} = B$, B est inversible et $B^{-1} = A$.

Théorème admis (il n'est pas évident que $AB = I_n$ implique $BA = I_n$).

Exemple d'utilisation de ce théorème. Soit A une matrice carrée telle que $A^2 - 5A + 6I_n = O_n$. Alors

$$A^2 - 5A = -6I_n \ ; \quad -\frac{1}{6}\left(A^2 - 5A\right) = I_n \ ; \ A\left[-\frac{1}{6}(A - 5I_n)\right] = I_n$$

Donc A est inversible, et $A^{-1} = -\frac{1}{6}(A - 5I_n)$.

Remarque. Si A, B vérifient l'égalité $AB = O_n$, la matrice A ne sera en général pas inversible.

Soit A telle que $A^2 - 3A = O_n$. Alors $A(A - 3I_n) = O_n$. Supposons A inversible ; alors, en multipliant l'égalité précédente par A^{-1}, on obtient $A = 3I_n$. Réciproquement, la matrice $3I_n$ est inversible. Dans tous les autres cas, une matrice vérifiant $A^2 - 3A = O_n$ n'est pas inversible.

Algorithme du pivot pour la recherche de A^{-1}

La matrice A est : inversible ssi on obtient par opérations élémentaires sur les lignes de A une matrice triangulaire sans zéros sur la diagonale ; non inversible ssi on obtient une matrice triangulaire avec un zéro sur la diagonale.

Si A est inversible, on effectue les mêmes opérations sur les matrices A et I_n, jusqu'à obtenir I_n et A^{-1} :

A	I_n	
…………		opérations élémentaires
I_n	A^{-1}	

Explication. Soit $A = \left(a_{i,j}\right)_{1 \leqslant i,j \leqslant 3}$ (on se place dans $\mathcal{M}_3(\mathbf{R})$ pour fixer les idées). D'après le théorème précédent, A est inversible ssi pour chacun des trois systèmes suivants il y a une solution unique :

$$\left\{\begin{array}{l} a_{1,1}x + a_{1,2}y + a_{1,3}z = 1 \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = 0 \\ a_{3,1}x + a_{3,2}y + a_{3,3}z = 0 \end{array}\right. , \quad \text{resp.} \begin{array}{c} 0 \\ 1 \\ 0 \end{array}, \quad \text{resp.} \begin{array}{c} 0 \\ 0 \\ 1 \end{array}$$

On utilise la méthode du pivot, en omettant les inconnues, et en écrivant donc uniquement la matrice A, puis la matrice unité pour les seconds membres. Voyons ceci sur un exemple :

$$A = \begin{pmatrix} \mathbf{1} & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{array}{l} \\ L_2 \leftarrow L_2 + L_1 \\ L_3 \leftarrow L_3 - L_1 \end{array}$$

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & \mathbf{2} & 0 \\ 0 & -2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad \begin{array}{l} L_1 \leftarrow 2L_1 - L_2 \\ \\ L_3 \leftarrow L_3 + L_2 \end{array}$$

L'opération élémentaire $L_3 \leftarrow L_3 + L_2$ permet d'obtenir 3 pivots non nuls ; il y a donc unicité de la solution pour chacun des trois membres, la matrice A est donc inversible. En vue d'obtenir A^{-1}, il vaut mieux effectuer simultanément $L_1 \leftarrow 2L_1 - L_2$.

$$\begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & \mathbf{2} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad L_1 \leftarrow L_1 + L_3$$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Les solutions des trois systèmes sont en évidence, elles forment la matrice de droite multipliée par $\frac{1}{2}$. On écrit

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

- Dans l'expression ci-dessus, gardez le 1/2 « en facteur ». Cela évite de trop nombreuses fractions et facilite les calculs ultérieurs.
- Vous adopterez la disposition pratique mise en évidence ci-dessus, en signalant que vous utilisez l'algorithme du pivot. Les opérations élémentaires sur la matrice A suffisent si on demande uniquement de déterminer si la matrice est inversible.

- En cas de « succès » (la matrice est inversible), vérifiez votre résultat en multipliant la matrice obtenue par la matrice de départ (on doit trouver la matrice unité). En cas d'erreur, le codage des opérations élémentaires devrait vous aider, il est donc capital d'y porter la plus grande attention.
- L'algorithme du pivot donne lieu à des calculs mécaniques, mais assez lourds, et ne devrait être utilisé qu'en dernier recours, si on n'a aucun renseignement sur la matrice étudiée (voir le théorème ci-dessus), et si l'énoncé ne suggère pas une autre méthode.

Application aux équations matricielles

Soit l'équation matricielle $AX = B$, avec $A \in \mathcal{M}_n(\mathbf{R})$.

Si A est inversible, $AX = B \Leftrightarrow X = A^{-1}B$.

En effet, $AX = B \Leftrightarrow A^{-1}(AX) = A^{-1}B$, et $\left(A^{-1}A\right)X = I_nX = X$.

X et B peuvent être deux matrices de $\mathcal{M}_n(\mathbf{R})$, ou deux matrices colonnes (voir ci-dessous)

Applications aux systèmes linéaires

Toujours pour fixer les idées, considérons le système linéaire à 3 inconnues :

$$(1)\left\{\begin{array}{l} a_{1,1}x + a_{1,2}y + a_{1,3}z = b_1 \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = b_2 \\ a_{3,1}x + a_{3,2}y + a_{3,3}z = b_3 \end{array}\right.$$

Ce système peut s'écrire $AX = Y$,

avec $A = \left(a_{i,j}\right)_{1 \leqslant i,j \leqslant 3}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $Y = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. Par conséquent :

Le système (1) est de Cramer (admet une solution unique) ssi la matrice $A = \left(a_{i,j}\right)_{1 \leqslant i,j \leqslant 3}$ est inversible, et on a alors $X = A^{-1}Y$.

On déduit de ceci une autre manière de déterminer si une matrice A donnée est inversible, et calculer le cas échéant son inverse : il suffit de résoudre le système (1). Si le système admet une solution unique (x, y, z), on a alors $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. Sinon, A n'est pas inversible.

Avec $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ dans $\mathcal{M}_2(\mathbf{R})$:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Leftrightarrow \begin{cases} 2x + y = b_1 \\ \ \ x + y = b_2 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = \ \ \ b_1 - \ \ b_2 \\ y = -b_1 + 2b_2 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

donc A est inversible, et $A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

Matrices triangulaires inversibles

Dans le cas d'un système triangulaire, on obtient (très important) :

Une matrice **triangulaire** est inversible ssi tous les éléments de sa diagonale sont non nuls.

N'écrivez jamais : cette matrice n'a pas de zéros sur la diagonale, elle est donc inversible. Utilisez uniquement cette propriété pour les matrices triangulaires, et précisez-le. La matrice $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ est inversible, avec uniquement des zéros sur la diagonale.

La matrice diagonale $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$ est inversible ssi tous les λ_i sont non nuls, et on a alors $D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & 0 \\ 0 & \frac{1}{\lambda_2} & 0 \\ 0 & 0 & \frac{1}{\lambda_3} \end{pmatrix}$ (évident).

2.4 Puissance n-ème d'une matrice carrée

Définition. Pour $A \in \mathcal{M}_n(\mathbf{R})$ et $n \in \mathbf{N}^*$, on définit $A^n = A \dots A$ (n facteurs). Si $A \neq \mathrm{O}_n$, on pose $A^0 = I_n$.

On a les règles de calcul :

$$A^m A^n = A^{m+n}\,;\ (A^m)^n = A^{mn}.$$

Mais attention, $(AB)^n$ n'est pas égal en général à A^nB^n, car la multiplication des matrices n'est pas commutative.

Pour une matrice diagonale $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$, on a, si $n \in \mathbf{N}^*$:

$D^n = \begin{pmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{pmatrix}$. Il n'y a pas de formule simple dès que la matrice n'est plus diagonale.

Pour calculer la puissance n-ème d'une matrice, on utilise souvent des propriétés de récurrence.

- La matrice $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ vérifie $B^2 = 3B$. On en déduit alors par récurrence : $\forall n \in \mathbf{N}^*, B^n = 3^{n-1}B$.

  Attention, la formule n'est pas valable pour $n = 0$. D'une manière générale, le calcul de A^0 demande une attention particulière.

- La matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ vérifie $A^3 = A^2 + 2A$. On montre alors, par récurrence : $\forall n \in \mathbf{N}^*, A^n = a_nA + b_nA^2$: la propriété est vraie pour $n = 1$, avec $a_1 = 1$, $b_1 = 0$, et si $A^n = a_nA + b_nA^2$, alors

  $$A^{n+1} = A^nA = \cdots = a_{n+1}A + b_{n+1}A^2,$$

  avec

  $$a_{n+1} = 2b_n,\ b_{n+1} = a_n + b_n,$$

  d'où la conclusion.

  Il reste à déterminer l'expression de a_n et b_n en fonction de n. Cela peut se faire à l'aide d'une suite linéaire récurrente à deux termes ; en effet on a, pour tout $n \in \mathbf{N}^*$:

  $$a_{n+2} = 2b_{n+1} = 2a_n + 2b_n = a_{n+1} + 2a_n$$

  Compte tenu de $a_1 = 1$, $b_1 = 0$, $a_2 = 0$, $b_2 = 1$, on trouve finalement

  $$a_n = \left(-\frac{2}{3}\right)(-1)^n + \left(\frac{1}{6}\right)2^n, \quad \text{puis} \quad b_n = \left(\frac{1}{3}\right)(-1)^n + \left(\frac{1}{6}\right)2^n.$$

Formule du binôme pour les matrices carrées

On démontre classiquement par récurrence :

> Soit A, B deux matrices de $\mathcal{M}_n(\mathbf{R})$ telles que $AB = BA$ (on dit alors que A et B **commutent**). Alors, pour tout $n \in \mathbf{N}^*$:
>
> $$(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^{n-k} B^k = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}$$

> N'oubliez pas de mentionner – et de vérifier – que les matrices commutent quand vous appliquez cette formule.

Cette formule est souvent utilisée avec une des deux matrices égale à kI_n, $k \in \mathbf{R}$, qui commute avec toute matrice de $\mathcal{M}_n(\mathbf{R})$.

- Soit à calculer A^n, avec $n \in \mathbf{N}^*$, et $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. On a calculé ci-dessus B^n, avec $B = A - I_3$, on a trouvé $B^n = 3^{n-1}B$ pour $n \geqslant 1$. $A = B + I_3$, donc $A^n = (B + I_3)^n$. B **et** I_3 **commutent**, donc d'après la formule du binôme :

  $$A^n = \sum_{k=0}^{n} \binom{n}{k} B^k {I_3}^{n-k} = \binom{n}{0} B^0 {I_3}^n + \sum_{k=1}^{n} \binom{n}{k} B^k {I_3}^{n-k}$$

  ($B^k = 3^{k-1}B$ est valable à partir de $k = 1$, il faut traiter à part $k = 0$.)

  $$A^n = I_3 + \sum_{k=1}^{n} \binom{n}{k} 3^{k-1} B = I_3 + \left(3^{-1} \sum_{k=1}^{n} \binom{n}{k} 3^k \right) B$$

  (Mise en facteur de B et de 3^{-1}. La somme est maintenant une somme de réels.)

  $$A^n = I_3 + \frac{1}{3} \left(\sum_{k=0}^{n} \binom{n}{k} 1^{n-k} 3^k - 1 \right) B = I_3 + \frac{1}{3} \left[(3+1)^n - 1 \right] B$$

  (On ajoute le terme manquant à la somme, de façon à pouvoir utiliser la formule du binôme pour les nombres réels, puis on le retire. Puis on utilise la formule du binôme.) On trouve finalement

  $$A^n = I_3 + \frac{1}{3}(4^n - 1) B$$

  On vérifie la formule avec $n = 1$. La formule est valable aussi avec $n = 0$.

- Soit à calculer A^n, avec $n \in \mathbf{N}^*$, et $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

  On décompose

  $A = I_3 + N$, avec $N = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. $N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $N^3 = O_3$,

  donc pour tout $k \geqslant 3$, $N^k = N^3 N^{k-3} = O_3$. La formule du binôme, applicable car I_3 et N commutent, fournit alors

  $$A^n = (I_3 + O_3)^n = \sum_{k=0}^{n} \binom{n}{k} {I_3}^{n-k} N^k$$
  $$= \binom{n}{0} N^0 + \binom{n}{1} N^1 + \binom{n}{2} N^2$$

  car les autres termes du développement sont nuls. Finalement :

  $$A^n = \begin{pmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}.$$

  On vérifie que ça marche avec $n = 1$. La formule est valable aussi avec $n = 0$.

Remarque. Une matrice N telle que $N^k = O_n$ pour quelque k est une matrice nilpotente. Une matrice triangulaire dont tous les éléments de la diagonale sont nuls est toujours une matrice nilpotente.

Application à l'étude de suites

Soit $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$, avec $n \in \mathbf{N}$. On suppose qu'il existe $A \in \mathcal{M}_3(\mathbf{R})$ telle que $\forall n \in \mathbf{N},\ X_{n+1} = AX_n$. On montre alors, par récurrence : $\forall n \in \mathbf{N},\ X_n = A^n X_0$. Et le calcul de A^n permet alors de déterminer l'expression de u_n, v_n, w_n pour tout $n \in \mathbf{N}$. (On a aussi, par récurrence : $X_n = A^{n-1} X_1$ si les suites sont définies sur $\mathbf{N}^*$.)

$$\begin{cases} u_{n+1} = 6u_n - v_n \\ v_{n+1} = u_n + 4v_n \end{cases} \Leftrightarrow X_{n+1} = AX_n \quad \text{avec}$$

$$A = \begin{pmatrix} 6 & -1 \\ 1 & 4 \end{pmatrix}, X_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}.$$

Par récurrence on a $X_n = A^n X_0$. $A^n = (5I_2 + J)^n$, avec

$$J = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix},$$

$J^k = \mathrm{O}_2$ pour tout $k \geqslant 2$. La formule du binôme fournit

$$A^n = 5^{n-1} \begin{pmatrix} 5+n & -n \\ n & 5-n \end{pmatrix},$$

puis

$$\begin{aligned} u_n &= 5^{n-1}\left((5+n)\,u_0 - nv_0\right) \\ v_n &= 5^{n-1}\left(nu_0 + (5-n)\,v_0\right) \end{aligned}$$

3. Un exemple d'espace vectoriel

3.1 Sous-espaces vectoriels, bases

$\mathcal{M}_{n,p}(\mathbf{R})$ est un espace vectoriel, voir § 2.2. On étudie ici comme première approche les espaces vectoriels $\mathcal{M}_{n,1}(\mathbf{R})$. On peut noter un élément de $\mathbf{R}^n$ sous forme d'une matrice colonne appartenant à $\mathcal{M}_{n,1}(\mathbf{R})$. On appellera « vecteur » un tel élément.

Soit le système homogène (3) étudié au § 1.3 :

$$(3)\begin{cases} a_{1,1}x + a_{1,2}y + a_{1,3}z = 0 \\ a_{2,1}x + a_{2,2}y + a_{2,3}z = 0 \\ a_{3,1}x + a_{3,2}y + a_{3,3}z = 0 \end{cases}$$

Ce système s'écrit sous forme matricielle $AX = \mathrm{O}$, avec

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix},\ X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ \mathrm{O} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

L'ensemble S des solutions de (3) vérifie les propriétés :

- $\mathrm{O} \in S$;
- si $X, Y \in S$, alors $X + Y \in S$;
- si $X \in S$ et $\alpha \in \mathbf{R}$, alors $\alpha X \in S$.

On dit alors que S est un **sous-espace vectoriel de $\mathcal{M}_{3,1}(\mathbf{R})$**.

Si S n'est pas réduit au vecteur nul O (ni étendu à $\mathcal{M}_{3,1}(\mathbf{R})$), on peut exprimer tout élément de S comme combinaison linéaire de p éléments fixés (appartenant à S), avec $p = 1$ ou $p = 2$.

Voyons ceci sur l'exemple du § 1.3. On rappelle les résultats obtenus :

$$(1)\begin{cases} -\lambda x + \ \ y + \ \ z = 0 \\ \ \ x - \lambda y + \ \ z = 0 \\ \ \ x + \ \ y - \lambda z = 0 \end{cases}.$$

Avec S_λ l'ensemble des solutions du système (1), on a :

$S_2 = \{z(1,1,1)\ ;\ z \in \mathbf{R}\}$;
$S_{-1} = \{y(-1,1,0) + z(-1,0,1)\ ;\ y, z \in \mathbf{R}\}$.
Dans les autres cas, $S_\lambda = \{(0,0,0)\}$.

- S_2 est l'ensemble des combinaisons linéaires d'un seul vecteur fixé, le vecteur X_1, avec $X_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. On dit que S_2 est le **sous-espace vectoriel de $\mathcal{M}_{3,1}(\mathbf{R})$ engendré par** (X_1). On note $S_2 = \text{Vect}(X_1)$. La famille (X_1) est une **base** de S_2 : tout élément de S_2 est combinaison linéaire des éléments de la base (ici, un seul élément) de façon unique.
- S_{-1} est l'ensemble des combinaisons linéaires de deux vecteurs X_2 et X_3 fixés, avec $X_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ et $X_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

  S_{-1} est le sous-espace vectoriel de $\mathcal{M}_{3,1}(\mathbf{R})$ engendré par (X_2, X_3).
  On note $S_{-1} = \text{Vect}(X_2, X_3)$.
  (X_2, X_3) est une base de S_{-1} : tout élément de S_{-1} est combinaison linéaire des deux éléments de la base de façon unique (on vérifie facilement que $xX_2 + yX_3 = x'X_2 + y'X_3 \Rightarrow x = x', y = y'$).
- Si $\lambda \notin \{-1, 2\}$, $S_\lambda = \{O\}$ est un sous-espace vectoriel de $\mathcal{M}_{3,1}(\mathbf{R})$, mais il n'a pas de base.

Les vecteurs $E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ constituent une base de $\mathcal{M}_{3,1}(\mathbf{R})$: tout vecteur $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathcal{M}_{3,1}(\mathbf{R})$ est combinaison linéaire des vecteurs E_1, E_2, E_3 de façon unique. Cette base est appelée **base canonique** de $\mathcal{M}_{3,1}(\mathbf{R})$, car les coordonnées du vecteur $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

dans cette base coïncident avec les coefficients de ce vecteur :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = xE_1 + yE_2 + zE_3$$

3.2 Applications linéaires

Définition

Une application $f : \mathcal{M}_{n,1}(\mathbf{R}) \to \mathcal{M}_{n,1}(\mathbf{R})$ est dite **linéaire** ssi :

$$\forall X, Y \in \mathcal{M}_{n,1}(\mathbf{R}), \quad f(X+Y) = f(X) + f(Y) \quad ;$$

$$\forall X \in \mathcal{M}_{n,1}(\mathbf{R}), \ \forall \alpha \in \mathbf{R}, \quad f(\alpha X) = \alpha f(X).$$

Avec $A \in \mathcal{M}_n(\mathbf{R})$, l'application

$$\mathcal{M}_{n,1}(\mathbf{R}) \to \mathcal{M}_{n,1}(\mathbf{R}) \ ; X \mapsto AX$$

est linéaire (évident d'après les règles de calcul sur les matrices), et on vérifie facilement que pour tout $j \in \{1, \ldots, n\}$, la j-ème colonne de la matrice A est égale à $f\left(E_j\right)$, où $(E_1, \ldots, E_n)$ est la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

Réciproquement, on a :

> Soit $f : \mathcal{M}_{n,1}(\mathbf{R}) \to \mathcal{M}_{n,1}(\mathbf{R})$ linéaire. Alors il existe une unique matrice carrée $A \in \mathcal{M}_{n,1}(\mathbf{R})$ telle que
>
> $$\forall X \in \mathcal{M}_{n,1}(\mathbf{R}), \quad f(X) = AX$$
>
> La matrice A est la matrice dont les colonnes successives sont les images $f(E_1), \ldots, f(E_n)$ des vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$.

Noyau, image d'une application linéaire

Soit $f : \mathcal{M}_{n,1}(\mathbf{R}) \to \mathcal{M}_{n,1}(\mathbf{R}), X \to AX$ une application linéaire.

Le **noyau** de f est l'ensemble noté $\mathrm{Ker}(f)$ des X appartenant à $\mathcal{M}_{n,1}(\mathbf{R})$ tels que $f(X) = \mathrm{O}$.

> $\mathrm{Ker}(f)$ est un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbf{R})$

En effet, $\mathrm{Ker}(f)$ est l'ensemble des solutions du système homogène $AX = \mathrm{O}$. On peut aussi le prouver directement :

- $\mathrm{O} \in \mathrm{Ker}(f)$, car $f(\mathrm{O}) = \mathrm{O}$;

- Si $X, X' \in \operatorname{Ker}(f)$, alors $f\left(X+X'\right)=f(X)+f\left(X'\right)$ car f est linéaire, donc $f\left(X+X'\right)=O$, donc $X+X' \in \operatorname{Ker}(f)$.
- Si $X \in \operatorname{Ker}(f)$ et $\alpha \in \mathbf{R}$, alors $f(\alpha X)=\alpha f(X)$ car f est linéaire, donc $f(\alpha X)=O$, donc $\alpha X \in \operatorname{Ker}(f)$.

L'**image** de f est l'ensemble $f\left(\mathcal{M}_{n,1}(\mathbf{R})\right)$. On le note $\operatorname{Im}(f)$:

$$\operatorname{Im}(f)=\left\{Y \in \mathcal{M}_{n,1}(\mathbf{R})\ ;\ \exists X \in \mathcal{M}_{n,1}(\mathbf{R}),\ \ f(X)=Y\right\}$$

$\operatorname{Im}(f)$ est un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbf{R})$

En effet, $O=f(O) \in \operatorname{Im}(f)$, et si $Y, Y' \in \operatorname{Im}(f)$ et $\alpha \in \mathbf{R}$, alors, d'après la linéarité de f :

$$Y+Y'=f(X)+f\left(X'\right)=f\left(X+X'\right) \in \operatorname{Im}(f)\ \ ;$$
$$\alpha Y=\alpha f(X)=f(\alpha X) \in \operatorname{Im}(f)\,.$$

Soit $X=\sum_{i=1}^{n} x_i E_i$ un élément quelconque de $\mathcal{M}_{n,1}(\mathbf{R})$, $(E_i)_{1\leqslant i\leqslant n}$ base canonique de $\mathcal{M}_{n,1}(\mathbf{R})$). On a alors

$$Y \in \operatorname{Im}(f) \Leftrightarrow Y=f(X) \Leftrightarrow Y=\sum_{i=1}^{n} x_i f(E_i)$$

$\operatorname{Im}(f)$ est donc l'ensemble des combinaisons linéaires des vecteurs $f(E_i)$. On a donc $\operatorname{Im}(f) = \operatorname{Vect}(f(E_1),\ldots,f(E_n))$, mais attention, $(f(E_1),\ldots,f(E_n))$ n'est pas nécessairement une base de $\operatorname{Im}(f)$.

Avec l'exemple du § 4.1.3, on a :

- $A_2=\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$; $\operatorname{Ker}(f_2)=S_2=\operatorname{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right)$.

  $\operatorname{Im}(f_2)=\operatorname{Vect}(F_1,F_2,F_3)$ avec F_1, F_2, F_3 les trois vecteurs colonnes de A_2. Pour $Y \in \operatorname{Im}(f_2)$, on a $Y=\gamma_1 F_1+\gamma_2 F_2+\gamma_3 F_3$, mais on a aussi $Y=(\gamma_1+1)F_1+(\gamma_2+1)F_2+(\gamma_3+1)F_3$, car $F_1+F_2+F_3=O$.
  Y est combinaison linéaire de F_1, F_2, F_3, mais pas de façon unique.
  Pour obtenir une base de $\operatorname{Im}(f_2)$, on remarque que
  $Y=\gamma_1 F_1+\gamma_2 F_2+\gamma_3(-F_1-F_2)=(\gamma_1-\gamma_2)F_1+(\gamma_2-\gamma_3)F_3$;
  $Y=z_1 F_1+z_2 F_2$: tout Y est combinaison linéaire de F_1 et F_2, et ceci de façon unique, comme on le vérifie sans peine. Ceci montre que (F_1, F_2) est une base de $\operatorname{Im}(f_2)$. On verra au chapitre 5 des méthodes moins acrobatiques pour déterminer une base d'un espace vectoriel ou d'un sous-espace vectoriel.

- $A_{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. ; $\text{Ker}\,(f_{-1}) = \text{S}_{-1} = \text{Vect}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right)$;

  $\text{Im}\,(f_{-1}) = \text{Vect}\,(G_1, G_1, G_1)$, avec $G_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. De façon évidente, (G_1, G_1, G_1) n'est pas une base de $\text{Im}\,(f_{-1})$, et (G_1) en est une.
- Dans les autres cas ($\lambda \notin \{-1\,;\,2\}$), $\text{Ker}\,(f_\lambda) = \{\text{O}\}$. Pour $\text{Im}\,(f)$: le système homogène (1) admet une solution unique (la solution O). Or on a vu que cette propriété (le système est de Cramer) dépendait uniquement des pivots du système, et pas des seconds membres (voir § 4.1.3). Cela veut dire que pour tout $(\alpha, \beta, \gamma) \in \mathbf{R}^3$, le système

$$\begin{cases} -\lambda x + y + z = \alpha \\ x - \lambda y + z = \beta \\ x + y - \lambda z = \gamma \end{cases}$$

  admet une solution unique, ou encore que la matrice A_λ est inver-sible, ou encore que, pour tout $Y = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$, il existe $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ unique tel que $f_\lambda\,(X) = Y$, c'est-à-dire que f_λ est une bijection de $\mathcal{M}_{3,1}\,(\mathbf{R})$ sur $\mathcal{M}_{3,1}\,(\mathbf{R})$. Par conséquent, $\text{Im}\,(f_\lambda) = \mathcal{M}_{3,1}\,(\mathbf{R})$.



Espaces vectoriels applications linéaires

5

1. Espaces vectoriels, sous-espaces vectoriels

1.1 Espaces vectoriels

Définition. Un **espace vectoriel sur R** (ev) est un ensemble E contenant au moins un élément, noté 0_E, ou simplement 0, muni d'une addition :

$$u \in E, \quad v \in E, \quad u + v \in E,$$

et d'une multiplication par les réels :

$$u \in E, \quad \alpha \in R, \quad \alpha \cdot u \in E,$$

avec les propriétés suivantes : $\forall u, v \in E$; $\forall \alpha, \beta \in \mathbf{R}$,

- $u + (v + w) = (u + v) + w$;
- $u + v = v + u$;
- $u + 0_E = 0_E + u = u$;
- $u + (-u) = (-u) + u = 0_E$ en notant $-u = (-1) \cdot u$;
- $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$;
- $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$;
- $\alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u$;
- $1 \cdot u = u$.

• Ces propriétés ne nécessitent aucun effort particulier de mémorisation et s'utilisent naturellement. La difficulté est le niveau d'abstraction inhabituel. Un **vecteur** est un élément d'un espace vectoriel : on sait additionner deux vecteurs, multiplier un vecteur par un nombre réel, avec toutes les « bonnes » propriétés. On peut bien sûr penser à l'ensemble des vecteurs du plan.

• $0 \cdot u = 0_E$, $\alpha \cdot 0_E = 0_E$, et $\alpha \cdot u = 0_E \Rightarrow \alpha = 0$ ou $u = 0_E$. Attention, quant A et B sont deux matrices carrées d'ordre n, le produit AB peut être nul sans que A ou B ne le soit.

Exemples. Les espaces vectoriels de référence sont les suivants :

• L'ensemble $\mathbf{R}^2 = \{(x, y) \,|x \in \mathbf{R}, y \in \mathbf{R}\}$ est un ev pour les opérations :

$$(x, y) + \left(x', y'\right) = \left(x + x', y + y'\right) \ ; \ \alpha (x, y) = (\alpha x, \alpha y)$$

On définit de même les ev $\mathbf{R}^n$ pour $n \in \mathbf{N}$ (avec $\mathbf{R}^1 = \mathbf{R}$, $\mathbf{R}^0 = \{0\}$).

• Avec n, p dans $\mathbf{N}^*$, l'ensemble $\mathcal{M}_{n,p}(\mathbf{R})$ des matrices réelles à n lignes et p colonnes est un ev pour l'addition des matrices et la multiplication d'une matrice par un nombre réel, voir § 4.2.2. Si E et F sont des ev, l'ensemble $\mathcal{L}(E, F)$ des applications linéaires de E dans F est un ev, voir § 5.3.1.

• L'ensemble $\mathbf{R}[X]$ des polynômes réels et les ensembles $\mathbf{R}_n[X]$ des polynômes réels de degré $\leqslant n$ ($n \in \mathbf{N}$) sont des espaces vectoriels pour les opérations $P + Q$, αP, voir le § 6.2 de l'Introduction. On a d'ailleurs les inclusions $\mathbf{R}_0[X] \subset \mathbf{R}_1[X] \subset \mathbf{R}_2[X] \subset \cdots \subset \mathbf{R}[X]$.

• L'ensemble $\mathbf{R}^{\mathbf{N}}$ des suites numériques est un ev pour les opérations :

$$(u_n)_{n\in\mathbf{N}} + (v_n)_{n\in\mathbf{N}} = (u_n + v_n)_{n\in\mathbf{N}} \ ; \ \alpha \cdot (u_n)_{n\in\mathbf{N}} = (\alpha u_n)_{n\in\mathbf{N}}$$

• Plus généralement, D étant un ensemble non vide, l'ensemble $\mathbf{R}^D$ des applications de D dans $\mathbf{R}$ est un ev pour les opérations

$$(f + g)(x) = f(x) + g(x) \ ; \ (\alpha \cdot f)(x) = \alpha f(x)$$

avec $f, g \in \mathbf{R}^D$; $\alpha \in \mathbf{R}$; $x \in D$.

Avec $D = \mathbf{N}$, on a bien l'ev des suites numériques. On retient (avec $D = I$) que l'ensemble des applications de l'intervalle réel non vide I dans $\mathbf{R}$ est un ev.

Avec $D = \Omega$, où $(\Omega, \mathcal{T}, \mathrm{P})$ est un espace probabilisé, on obtient que l'ensemble des variables aléatoires définies sur Ω est un ev.

1.2 Sous-espaces vectoriels

Définition. Soit E un ev. On dit que l'ensemble F est un **sous-espace vectoriel** de E ssi F est un ev et $F \subset E$. Abrégé : F est un sev de E.

• $\{0_E\}$ et E sont des sev de l'ev E.

• $\mathbf{R}_0[X]$ (ev des fonctions constantes) est un sev de $\mathbf{R}_1[X]$ (ev des fonctions affines), lui-même sev de $\mathbf{R}_2[X]$... Et, pour tout $n \in \mathbf{N}$, l'ev $\mathbf{R}_n[X]$ des polynômes de degré $\leqslant n$ est un sev de $\mathbf{R}[X]$.

Théorème. F est un sev de E ssi :

• $F \subset E$ et $0_E \in F$;

• $u \in F, v \in F \Rightarrow u + v \in F$;

• $u \in F, \alpha \in \mathbf{R} \Rightarrow \alpha \cdot u \in F$.

On utilise ce théorème — et jamais la définition — pour montrer qu'un ensemble est un ev, en établissant que l'ensemble en question est un sev d'un ev de référence.

• Pour $n \in \mathbf{N} \cup \{\infty\}$, l'ensemble $\mathrm{C}^n(I)$ des applications de classe C^n de l'intervalle I dans $\mathbf{R}$ est un ev. En effet, d'après § 1.3.1 :

– $\mathrm{C}^n(I)$ est inclus dans l'espace vectoriel de référence $\mathbf{R}^I$, et la fonction nulle $I \to \mathbf{R}$, $x \mapsto 0$ appartient à $\mathrm{C}^n(I)$;

– si $f, g \in \mathrm{C}^n(I)$, alors $f + g \in \mathrm{C}^n(I)$;

– si $f \in \mathrm{C}^n(I)$ et $\alpha \in \mathbf{R}$, alors $\alpha f \in \mathrm{C}^n(I)$.

$\mathrm{C}^n(I)$ est donc un sev de $\mathbf{R}^D$, donc un ev. On peut préciser que si $n \geqslant n'$, $\mathrm{C}^n(I)$ est un sev de $\mathrm{C}^{n'}(I)$.

• L'ensemble $F = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \; ; \; a + d = 0 \right\}$ est un ev car c'est un sev de $\mathcal{M}_2(\mathbf{R})$. En effet :

– $F \subset \mathcal{M}_2(\mathbf{R})$ et la matrice nulle $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ appartient à F ;

– si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et $M' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ appartiennent à F et α à $\mathbf{R}$, alors $M + M' = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$ et $\alpha M = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix}$ appartiennent à F car

$$\left(a + a'\right) + \left(d + d'\right) = (a + d) + \left(a' + d'\right) = 0 + 0 = 0$$

et

$$\alpha a + \alpha d = \alpha (a + d) = \alpha 0 = 0.$$

• L'ensemble $G = \{u_n \mid \forall n \in \mathbf{N}, u_{n+2} = -2u_{n+1} + 3u_n\}$ est un sev de l'ev $\mathbf{R}^{\mathbf{N}}$ des suites numériques, en effet :

– $G \subset \mathbf{R}^{\mathbf{N}}$;

– la suite nulle ($\forall n \in \mathbf{N}, u_n = 0$) appartient à G ;

– si $(u_n), (v_n)$ sont deux suites appartenant à G et α un nombre réel, alors $(u_n) + (v_n) = (u_n + v_n)$ et $\alpha \cdot (u_n) = (\alpha u_n)$ appartiennent à G. En effet, pour tout n de $\mathbf{N}$:

$$\begin{aligned} u_{n+2} + v_{n+2} &= -2u_{n+1} + 3u_n + (-2v_{n+1} + 3v_n) \\ &= -2(u_{n+1} + v_{n+1}) + (u_n + v_n) \end{aligned}$$

$$\alpha u_{n+2} = \alpha(-2u_{n+1} + 3u_n) = -2(\alpha u_{n+1}) + 3(\alpha u_n)$$

Notion de combinaison linéaire. Soient $u_1, \dots, u_p$ des éléments de l'ev E, et $\alpha_1, \alpha_2, \dots, \alpha_p$ des nombres réels. Le vecteur

$$\sum_{i=1}^{p} \alpha_i u_i = \alpha_1 u_1 + \cdots + \alpha_p u_p$$

est une **combinaison linéaire** des vecteurs $u_1, \dots, u_p$.

L'ensemble de toutes les combinaisons linéaires des vecteurs fixés $u_1, \dots, u_p$ de E est un sev de E, appelé **sous-espace vectoriel engendré par** $u_1, \dots, u_p$, et noté $\mathrm{Vect}(u_1, \dots, u_p)$:

$$\mathrm{Vect}(u_1, \dots, u_p) = \left\{ u \in E \mid \exists (\alpha_1, \dots, \alpha_p) \in \mathbf{R}^p, u = \sum_{i=1}^{p} \alpha_i u_i \right\}$$

• On a $\mathrm{Vect}(0_E) = \{0_E\}$.

• $F = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} ; a, b \in \mathbf{R} \right\}$ est un sev de $\mathcal{M}_2(\mathbf{R})$, donc un ev, car

$$F = \mathrm{Vect}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right)$$

Notez que les vecteurs $u_1, \dots, u_p$ appartiennent à $\mathrm{Vect}(u_1, \dots, u_p)$:

$$u_1 = 1 \cdot u_1 + 0 \cdot u_2 + \cdots + 0 \cdot u_p \quad \text{par exemple.}$$

0_E appartient bien sûr à $\mathrm{Vect}(u_1, \dots, u_p)$: $\quad 0_E = 0 \cdot u_1 + \cdots + 0 \cdot u_p$

1.3 Familles libres, génératrices, bases

Définitions. Soit $p \in \mathbf{N}^*$, E un ev et $\mathcal{F} = \left(u_1, \ldots, u_p\right)$ une famille de vecteurs de E. On dit que la famille $\mathcal{F}$ est :

• **génératrice de** E ssi tout vecteur de E est combinaison linéaire des éléments de $\mathcal{F}$:

$$\forall u \in E, \quad \exists \left(\alpha_1, \ldots, \alpha_p\right) \in \mathbf{R}^p, \quad u = \sum_{i=1}^{p} \alpha_i u_i$$

On dit aussi que $\mathcal{F}$ **engendre** E ;

• **libre** ssi le vecteur nul 0_E est combinaison linéaire des éléments de $\mathcal{F}$ de façon unique :

$$\alpha_1 u_1 + \cdots + \alpha_p u_p = 0_E \Rightarrow \alpha_1 = \cdots = \alpha_p = 0$$

• une **base de** E ssi tout vecteur de E est combinaison linéaire des éléments de $\mathcal{F}$ de façon unique :

$$\forall u \in E,\ \exists \left(\alpha_1, \ldots, \alpha_p\right) \in \mathbf{R}^p,\ \left(\ \alpha_1, \ldots, \alpha_p\right) \text{ unique},\ u = \sum_{i=1}^{p} \alpha_i u_i$$

Les réels $\alpha_1, \ldots, \alpha_p$ sont alors les **coordonnées** du vecteur u dans la base $\mathcal{F}$, et on dit que E est de **dimension finie**.

Théorème de la dimension. Dans un espace vectoriel de dimension finie E, toutes les bases ont le même nombre d'éléments. Ce nombre noté $\dim(E)$ est appelé la **dimension** de E.

Soit $\mathcal{F}$ une famille d'éléments de E de dimension finie n. Les propriétés suivantes sont équivalentes :

• $\mathcal{F}$ est une base de E ;
• $\mathcal{F}$ est libre et de cardinal n ;
• $\mathcal{F}$ est génératrice de E et de cardinal n ;
• $\mathcal{F}$ est libre et génératrice de E.

On utilise ce théorème principalement pour montrer qu'une famille $\mathcal{F}$ est une base de E. On utilisera surtout (avec E de dimension n) :

libre et de cardinal $n \Rightarrow$ base
libre et génératrice de $E \Rightarrow$ base

• $((1\,;\,0)\,,(0\,;\,1))$ est une base de $\mathbf{R}^2$, qui est donc de dimension 2 (voir plus loin). Soit $\mathcal{B} = (u_1, u_2)$, avec $u_1 = (2\,;\,-1)\,, u_2 = (1\,;\,2)$.

$\mathcal{B}$ est une famille libre, de cardinal 2, donc $\mathcal{B}$ est une base de $\mathbf{R}^2$.

• Soit $F = \{(x\,;y\,;z) \in \mathbf{R}^3 | x - y + z = 0\}$.

$u = (x\,;y\,;z) \in F \Leftrightarrow x = y - z \Leftrightarrow u = (y - z\,;y\,;z) \Leftrightarrow u = yu_1 + zu_2$, avec $u_1 = (1\,;1\,;0)\,, u_2 = (-1\,;0\,;1)$.

F est donc le sous-espace vectoriel de $\mathbf{R}^2$ engendré par la famille de vecteurs $\mathcal{B} = (u_1\ ;\ u_2)$; donc F est un espace vectoriel, dont une famille génératrice est $\mathcal{B}$. Comme cette famille est libre, $\mathcal{B}$ est une base de l'espace vectoriel F (qui est donc de dimension 2).

Quand une famille est libre, on dit que les vecteurs qui la composent sont **linéairement indépendants**. Une **famille liée** est une famille non libre. À cause du théorème de la dimension, il est important de reconnaître les familles libres. Dans ce sens on a :

• (u) est une famille libre ssi $u \neq 0_E$.

• Les vecteurs u, v de l'ev E sont dits **colinéaires** ssi $u = 0_E$ ou $v = \lambda u$, avec $\lambda \in \mathbf{R}$. (u, v) est une famille libre ssi u et v ne sont pas colinéaires.

Soient $u = (2\,;\,-1\,;3)\,, v = (-4\,;2\,;\,-6)\,, w = (-4\,;2\,;6)$. u, v sont colinéaires ($v = -2u$), ils ne forment donc pas une famille libre. (u, w) et (v, w) sont des familles libres.

• Soient $\mathcal{F} = (u, v, w)$ une famille de trois vecteurs. Si deux d'entre eux sont colinéaires, alors la famille $\mathcal{F}$ est liée, mais la réciproque est fausse.

$u = (1\,;0\,;\,-1)\,, v = (2\,;3\,;5)\,, w = (-1\,;0\,;1)$: La famille $\mathcal{F}$ est liée car les vecteurs u et w sont colinéaires.

$u = (1\,;1\,;\,-1)\,, v = (2\,;\,-1\,;2)\,, w = (3\,;0\,;1)$. $\mathcal{F}$ est une famille liée car $u + v - w = 0$, alors que les vecteurs u, v, w ne sont pas deux à deux colinéaires.

Pour montrer qu'une famille de plus de deux vecteurs est libre, on sera amené (très souvent) à résoudre le système linéaire correspondant, qui est un système homogène : la famille est libre ssi le système admet uniquement la solution nulle.

Exemples de bases et de dimensions

• Avec $n \in \mathbf{N}^*$, $\mathbf{R^n}$ est un espace vectoriel de dimension n. Par exemple, $\mathbf{R}^3$ est de dimension 3. Soit $\mathcal{B} = (e_1, e_2, e_3)$, avec

$$e_1 = (1\,;0\,;0)\,,\ e_2 = (0\,;1\,;0)\,,\ e_3 = (0\,;0\,;1)$$

$\mathcal{B}$ est une base de $\mathbf{R}^3$, appelée la **base canonique** de $\mathbf{R}^3$.

• L'espace vectoriel $\{0\}$ n'a pas de base, il est de dimension 0.

• Soit E un ev dimension n. Le seul sev de E de dimension 0 est $\{0_E\}$, le seul sev de dimension n est E.

• Les sev de $\mathbf{R}^3$ de dimension 1 sont les $\text{Vect}(u) = \{x \cdot u \mid x \in \mathbf{R}\}$ avec $u \neq 0$ (droites vectorielles), les sev de dimension 2 sont les $\text{Vect}(u, v) = \{x \cdot u + y \cdot v \mid x, y \in \mathbf{R}\}$ avec u, v non colinéaires (plans vectoriels).

• Avec $n \in \mathbf{N}$, l'espace vectoriel $\mathbf{R}_n[X]$ des polynômes de degré $\leqslant n$ est de dimension $n+1$. $\mathcal{B} = \left(1, X, X^2\right)$ est la base canonique de $\mathbf{R}_2[X]$, car pour tout $P \in \mathbf{R}_2[X]$, $P(X) = a_0 + a_1 X + a_2 X^2$ de façon unique.

• Avec $n, p \in \mathbf{N}^*$, $\mathcal{M}_{n,p}(\mathbf{R})$ est un espace vectoriel de dimension np. $\mathcal{M}_n(\mathbf{R})$ est de dimension n^2. La base canonique de $\mathcal{M}_2(\mathbf{R})$ est

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

• Pour montrer que (u, v, w) est une base de $\mathbf{R}^3$, la démarche courante, illustrée par un exemple, est la suivante :

Avec $u = (1\,;1\,;\,-1)\,, v = (-1\,;1\,;1)\,, w = (-1\,;1\,;1)$:

$$x \cdot u + y \cdot v + z \cdot w = (0\,;0\,;0) \Leftrightarrow \begin{cases} x - y - z = 0 \\ x + y + z = 0 \\ -x + y + z = 0 \end{cases} \Leftrightarrow \cdots \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases}$$

La famille (u, v, w) est donc libre, et de cardinal 3 ; $\mathbf{R}^3$ est de dimension 3, il en résulte que (u, v, w) est une base de $\mathbf{R}^3$.

On utilisera cette méthode quand on n'a aucun renseignement sur la famille (u, v, w) ; le résultat peut provenir d'autres considérations (matrices inversibles, théorie du changement de base, endomorphismes diagonalisables, voir plus loin).

• Ne confondez pas dimension et cardinal : dans un espace vectoriel de dimension n, toutes les bases ont le même cardinal, mais ne parlez pas de cardinal d'un espace vectoriel, ni de dimension d'une base.

• Dans un ev E de dimension n, une famille libre a au plus n éléments. Si elle a moins de n éléments, on peut la compléter de façon à obtenir une base (théorème de la base incomplète). Si elle a exactement n éléments, c'est une base de E.

• Dans un ev E de dimension n, une famille génératrice a au moins n éléments. Si elle a plus de n éléments, on peut en extraire une sous-famille libre de cardinal n (ou de cardinal maximal si n n'est pas connu), qui est alors une base de E. Si elle a exactement n éléments, c'est une base de E.

2. Applications linéaires

2.1 Définition, exemples

Définitions. Soit E, F deux ev. Une application $f : E \to F$ est dite **linéaire** ssi :

- $\forall u, v \in E, f(u+v) = f(u) + f(v)$;
- $\forall u \in E, \forall \alpha \in \mathbf{R}, f(\alpha \cdot u) = \alpha \cdot f(u)$.

Soit $f : E \to F$ linéaire.

- Si f est bijective, on dit que f est un **isomorphisme** de E sur F.
- Si $F = E$, on dit que $f : E \to E$ est un **endomorphisme** de E.
- Si $f : E \to E$ est bijective, on dit que f est un **automorphisme** de E.
- Si $F = \mathbf{R}$, on dit que $f : E \to \mathbf{R}$ est une **forme linéaire**.

Pour prouver que $f : E \to F$ est linéaire, on utilise la définition. Pour prouver que f linéaire est : un isomorphisme, il suffit de prouver qu'elle est bijective ; un endomorphisme : $f(E) \subset E$; un automorphisme : $f(E) \subset E$ et f bijective.

Premières propriétés. Pour $f : E \to F$ linéaire :

- $f(0_E) = 0_F$;
- $\forall u \in E,\ f(-u) = -f(u)$;
- $\forall u_1, \ldots, u_p \in E,\ f\left(\sum_{i=1}^{p} \alpha_i u_i\right) = \sum_{i=1}^{p} \alpha_i f(u_i)$

Cette dernière propriété est une propriété caractéristique des applications linéaires : f est une application linéaire ssi l'image par f d'une combinaison linéaire est la combinaison linéaire des images.

Premiers exemples

- L'application nulle $E \to F, u \mapsto 0_F$ est linéaire.
- L'application identique $\mathrm{Id}_E : E \to E, u \mapsto u$ est linéaire. C'est d'ailleurs un automorphisme de E.
- La transposée d'une matrice carrée (voir § 4.2.1) définit une endomorphisme de $\mathcal{M}_n(\mathbf{R})$:

$$ {}^t\left(M + M'\right) = {}^tM + {}^tM' \ ;\ {}^t(\alpha M) = \alpha\, {}^tM. $$

- Les applications linéaires de $\mathbf{R}$ dans $\mathbf{R}$ sont les applications $x \mapsto ax$.
- Les formes linéaires de $\mathbf{R}^3$ dans $\mathbf{R}$ sont les applications

$$ (x, y, z) \mapsto ax + by + cz $$

2.2 Applications linéaires et matrices

Soit $f : E \to F$ linéaire, $\mathcal{B}$ une base de E. **L'image par f du vecteur $u \in E$ est entièrement déterminée par la donnée des images $f(e_1), \ldots, f(e_p)$ des vecteurs de la base $\mathcal{B}$.**

Plus précisément, le théorème suivant permet de calculer effectivement sur les applications linéaires :

> **Théorème.** Soit E, F deux ev ; $\mathcal{B} = (e_1, \ldots, e_p)$ une base de E ; $\mathcal{C} = (e'_1, \ldots, e'_n)$ une base de F ; f une application linéaire de E dans F.
>
> Soit :
>
> • $Y = \mathrm{Mat}(f(u), \mathcal{C})$ la matrice colonne des coordonnées de $f(u)$ dans la base $\mathcal{C}$;
>
> • $M = \mathrm{Mat}(f, \mathcal{B}, \mathcal{C})$ la matrice dont les colonnes successives sont les coordonnées des vecteurs $f(e_1), \ldots, f(e_p)$ dans la base $\mathcal{C}$;
>
> • $X = \mathrm{Mat}(u, \mathcal{B})$ la matrice colonne des coordonnées de u dans la base $\mathcal{B}$.
>
> Alors
> $$Y = MX$$

En effet, si $u = \sum_{i=1}^{p} x_i e_i$, alors

$$f(u) = f\left(\sum_{i=1}^{p} x_i e_i\right) = \sum_{i=1}^{p} x_i f(e_i) = \sum_{i=1}^{p} x_i \left(\sum_{j=1}^{n} a_{i,j} e'_j\right)$$
$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{p} a_{i,j} x_j\right) e'_j$$

d'où la conclusion, avec $M = (a_{i,j})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$.

• La matrice de l'application linéaire nulle $E \to F, u \mapsto 0_F$, est toujours la matrice nulle de $\mathcal{M}_{n,p}(\mathbf{R})$.

• Soient $\mathcal{B} = (e_1, e_2, e_3)$ et $\mathcal{C} = (e'_1, e'_2)$ les bases canoniques de $\mathbf{R}^3$ et $\mathbf{R}^2$, et $f : \mathbf{R}^3 \to \mathbf{R}^2$ linéaire telle que

$$f(e_1) = (1\,;2)\,,\ f(e_2) = (3\,;4)\,,\ f(e_3) = (5\,;6)\,.$$

$$f(e_1) = 1 \cdot e'_1 + 2 \cdot e'_2, \ldots \text{ donc } M = \mathrm{Mat}(f, \mathcal{B}, \mathcal{C}) = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Avec $u = xe_1 + ye_2 + ze_3$, on a $X = \mathrm{Mat}(u, \mathcal{B}) = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

$$Y = MX = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 3y + 5z \\ 2x + 4y + 6z \end{pmatrix}$$

On a donc $f(u) = (x + 3y + 5z, 2x + 4y + 6z)$, ce que confirme le calcul direct $f(u) = xf(e_1) + yf(e_2) + zf(e_3) = \ldots$

• Soit la forme linéaire $f : \mathbf{R}^3 \to \mathbf{R}, (x, y, z) \mapsto x + 2y + 3z$.

Avec $\mathcal{B} = (e_1, e_2, e_3)$ base canonique de $\mathbf{R}^3$, $\mathcal{C} = (1)$ base canonique de $\mathbf{R}$, on a $f(e_1) = 1, f(e_2) = 2, f(e_3) = 3$, donc

$$\mathrm{Mat}(f, \mathcal{B}, \mathcal{C}) = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \quad ; \quad f(x\,;y\,;z) = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

en identifiant $\mathcal{M}_{3,1}(\mathbf{R})$ et $\mathbf{R}^3$, $\mathcal{M}_{1,1}(\mathbf{R})$ et $\mathbf{R}$.

Cas des endomorphismes

Dans le cas, le plus fréquent, où f est un endomorphisme de E, on note simplement $\mathrm{Mat}(f, \mathcal{B})$ au lieu de $\mathrm{Mat}(f, \mathcal{B}, \mathcal{B})$. Cette matrice est appelée **matrice de** f **dans la base** $\mathcal{B}$, ou **relativement à la base** $\mathcal{B}$.

De même, la matrice colonne $\mathrm{Mat}(u, \mathcal{B})$ est appelée matrice de u dans la base $\mathcal{B}$.

Pour f endomorphisme de E, avec $\mathcal{B}$ base de E :

$$\mathrm{Mat}(f(u), \mathcal{B}) = \mathrm{Mat}(f, \mathcal{B}) \times \mathrm{Mat}(u, \mathcal{B})$$

Ainsi, dans toute base $\mathcal{B}$ d'un espace vectoriel E de dimension n, on a $\mathrm{Mat}(\mathrm{Id}_E, \mathcal{B}) = I_n$. Si $X = \mathrm{Mat}(u, \mathcal{B})$, alors $I_n X = X$, ce qui correspond bien à $\mathrm{Id}_E(u) = u$.

Ne pas confondre le vecteur $u \in E$ (qui peut être un polynôme, une fonction, une matrice...) avec la matrice colonne $\mathrm{Mat}(u, \mathcal{B})$ des coordonnées de u dans la base $\mathcal{B}$, sauf si E est l'espace vectoriel $\mathbf{R}^n$ et $\mathcal{B}$ sa base canonique.

• Le polynôme $P = a + bX + cX^2$ a pour coordonnées (a, b, c) dans la base canonique $\mathcal{B} = (1, X, X^2)$ de $\mathbf{R}_2[X]$, mais n'est pas égal au vecteur (a, b, c) de $\mathbf{R}^3$.

• En revenant à la définition, on montre que $f : P \mapsto f(P) = P + P'$ est un endomorphisme de $\mathbf{R}_2[X]$, dont la matrice dans la base $\mathcal{B}$ est $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Le calcul $A \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ montre que l'image du polynôme $2 - 2X + X^2$ par f est le polynôme X^2, ce que confirme le calcul direct.

• Soit g l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique est la matrice A précédente. Le même calcul s'interprète maintenant en disant que $g((2, -2, 1)) = (0, 0, 1)$.

• Soit $k \in \mathbf{R}$. L'application $E \to E\,;\, u \mapsto k \cdot u$ est un endomorphisme dont la matrice est kI_n dans toute base de E, mais en dehors de ce cas, un endomorphisme est représenté par une matrice différente quand on change de base, voir § 6.1, et l'exemple ci-dessous.

Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ est $A = \mathrm{Mat}(f, \mathcal{B}) = \begin{pmatrix} 2 & -2 & -2 \\ -1 & 3 & 1 \\ -1 & -1 & 1 \end{pmatrix}$.

Soit $\mathcal{B}' = (u_1, u_2, u_3)$, avec

$$u_1 = e_1 - e_2 \;;\; u_2 = e_1 + e_3 \;;\; u_3 = e_2 - e_3$$

$\mathcal{B}'$ est une base de $\mathbf{R}^3$, car c'est une famille libre de cardinal 3. On calcule $f(u_1)$ en utilisant $\mathrm{Mat}(f, \mathcal{B}) \times \mathrm{Mat}(u_1, \mathcal{B}) = \mathrm{Mat}(f(u_1), \mathcal{B})$:

$$\begin{pmatrix} 2 & -2 & -2 \\ -1 & 3 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \\ 0 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Donc $f(u_1) = (4, -4, 0) = 4u_1 = 4 \cdot u_1 + 0 \cdot u_2 + 0 \cdot u_3$

On trouve de même

$$f(u_2) = (0, 0, 0) = 0 \cdot u_1 + 0 \cdot u_2 + 0 \cdot u_3$$

$$f(u_3) = (0, 2, -2) = 2u_3 = 0 \cdot u_1 + 0 \cdot u_2 + 2 \cdot u_3$$

Par définition de la matrice de f dans la base $\mathcal{B}'$, on a donc :

$$A' = \mathrm{Mat}(f, \mathcal{B}') = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \neq A = \mathrm{Mat}(f, \mathcal{B})$$

3. Espace vectoriel $\mathcal{L}(E, F)$, algèbre $\mathcal{L}(E)$

3.1 Espace vectoriel $\mathcal{L}(E, F)$

Théorème 1. Soit E et F deux ev de dimension finie, $\mathcal{B}$ une base de E, $\mathcal{C}$ une base de F.

• Alors l'ensemble $\mathcal{L}(E, F)$ des applications linéaires de E dans F est un espace vectoriel pour les opérations $f + g$ et αf.

• L'application Φ définie par

$$\Phi : \mathcal{L}(E, F) \to \mathcal{M}_{n,p}(\mathbf{R}) \;\; ; \; f \mapsto A = \text{Mat}(f, \mathcal{B}, \mathcal{C})$$

est un isomorphisme, c'est-à-dire :

– Φ est une bijection :

$$\forall A \in \mathcal{M}_{n,p}(\mathbf{R}), \exists f \in \mathcal{L}(E, F), f \text{ unique}, \text{Mat}(f, \mathcal{B}, \mathcal{C}) = A$$

– Φ est linéaire : $\forall f, g \in \mathcal{L}(E, F), \forall \alpha \in \mathbf{R}$,

$$\text{Mat}(f + g, \mathcal{B}, \mathcal{C}) = \text{Mat}(f, \mathcal{B}, \mathcal{C}) + \text{Mat}(g, \mathcal{B}, \mathcal{C})$$

$$\text{Mat}(\alpha f, \mathcal{B}, \mathcal{C}) = \alpha \, \text{Mat}(f, \mathcal{B}, \mathcal{C})$$

Attention, la matrice associée à une application linéaire n'est pas unique, elle dépend des bases choisies dans les espaces E et F. Si $E = F = \mathbf{R}^n$ (très fréquent), on donne naturellement la matrice de f dans la base canonique de $\mathbf{R}^n$, mais il peut exister des bases de $\mathbf{R}^n$ dans la quelle la matrice de f sera plus simple, voir chapitre 6.

3.2 Algèbre $\mathcal{L}(E)$

L'espace vectoriel $\mathcal{L}(E, E)$ des endomorphismes de E est noté simplement $\mathcal{L}(E)$. Soit $\mathcal{B}$ une base de E, de dimension n. D'après ce qui précède :

• $\forall f \in \mathcal{L}(E)$, $\exists A \in \mathcal{M}_n(\mathbf{R})$, A unique, $\text{Mat}(f, \mathcal{B}) = A$

• $\forall f, g \in \mathcal{L}(E)$, $\forall \alpha \in \mathbf{R}$,

$$\text{Mat}(f + g, \mathcal{B}) = \text{Mat}(f, \mathcal{B}) + \text{Mat}(g, \mathcal{B})$$

$$\text{Mat}(\alpha f, \mathcal{B}) = \alpha \text{Mat}(f, \mathcal{B})$$

On a de plus le

> **Théorème 2**
>
> • Si $f, g \in \mathcal{L}(E)$, alors $f \circ g \in \mathcal{L}(E, F)$, et
>
> $$\text{Mat}(f \circ g, \mathcal{B}) = \text{Mat}(f, \mathcal{B}) \times \text{Mat}(g, \mathcal{B})$$
>
> • $\text{Mat}(\text{Id}_E, \mathcal{B}) = I_n$
>
> • $f \in \mathcal{L}(E, F)$ est bijective ssi $A = \text{Mat}(f, \mathcal{B})$ est inversible, et on a alors :
>
> $$\text{Mat}\left(f^{-1}, \mathcal{B}\right) = A^{-1} = [\text{Mat}(f, \mathcal{B})]^{-1}$$

Remarques

• Ces résultats sont cohérents avec les propriétés des bijections réciproques. Si $f : E \to E$ est bijective, on sait que $f \circ f^{-1} = f^{-1} \circ f = \text{Id}_E$, ce qui correspond à, avec $A = \text{Mat}(f, \mathcal{B})$:

$$A \times A^{-1} = A^{-1} \times A = I_n$$

• Soit $Aut(E)$ l'ensemble des automorphismes de E. On a les propriétés : $\text{Id}_E \in \text{Aut}(E)$; si $f, g \in \text{Aut}(E)$, alors $f \circ g \in \text{Aut}(E)$, et $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. La composition des automorphismes est associative $((f \circ (g \circ h) = (f \circ g) \circ h)$, mais pas commutative ($f \circ g \neq g \circ f$ en général). On résume ces propriétés en disant que $\text{Aut}(E)$ est un groupe (non commutatif) pour la composition des applications. $\text{Aut}(E)$ est aussi noté $\text{GL}(E)$ (« groupe linéaire de E »).

Une application

Puissance n-ème d'un endomorphisme. Soit $f \in \mathcal{L}(E)$, et soit $p \in \mathbf{N}^*$. On pose

$$f^p = f \circ \cdots \circ f (p \text{ termes})$$

Si f n'est pas l'endomorphisme nul, on pose $f^0 = \text{Id}_E$. On a alors :

$$\text{Mat}(f^p, \mathcal{B}) = [\text{Mat}(f, \mathcal{B})]^p$$

Remarque. Soit $A = \text{Mat}(f, \mathcal{B})$. Si f est bijective (ce qui est équivalent à A inversible), on peut définir A^k et f^k avec k entier négatif. En effet, si A est inversible et $p \in \mathbf{N}$, alors A^p est inversible, et $(A^p)^{-1} = \left(A^{-1}\right)^p$ (car $A^p \times \left(A^{-1}\right)^p = I_n$). En posant

$$A^{-p} = (A^p)^{-1} = \left(A^{-1}\right)^p \; ; \; f^{-p} = (f^p)^{-1} = \left(f^{-1}\right)^p,$$

on a alors, maintenant pour tout $p \in \mathbf{Z}$: $\text{Mat}(f^p, \mathcal{B}) = [\text{Mat}(f, \mathcal{B})]^p$.

Mais attention, ceci est valable uniquement si f est bijective.

Dans un contexte différent (si f est une fonction de $\mathbf{R}$ dans $\mathbf{R}$), f^n désigne la fonction $x \mapsto (f(x))^n$. Mais aucune ambiguïté n'est possible, le produit de deux vecteurs n'étant pas une opération définie.

4. Noyau et image d'une application linéaire

4.1 Définitions, propriétés

Définitions. Soit E, F deux ev de dimension finie, et $f : E \to F$ linéaire.

Le **noyau** de f est le sous-ensemble de E, noté Ker (f), et défini par

$$\operatorname{Ker}(f) = \{u | u \in E \ ; \ f(u) = 0_F\}$$

L'**image** de f est le sous-ensemble de F, noté Im (f), et défini par

$$\operatorname{Im}(f) = \{v | v \in F \, ; \exists u \in E, f(u) = v\}$$

$u \in \operatorname{Ker}(f) \Leftrightarrow f(u) = 0_F$. La détermination de Ker (f) conduit donc naturellement à la résolution d'un système linéaire homogène (voir § 4.1.3).

Pour la détermination de Im (f), voir ci-dessous.

Propriétés

- Ker (f) est un sev de E.
- f injective $\Leftrightarrow \operatorname{Ker}(f) = \{0_E\}$
- Im (f) est un sev de F.
- f surjective $\Leftrightarrow \operatorname{Im}(f) = F$
- Si $(e_1, \ldots, e_p)$ est une base de E, alors la famille $(f(e_1), \ldots, f(e_p))$ est une famille génératrice de Im (f).
- (**formule du rang**) $\dim(\operatorname{Ker}(f)) + \dim(\operatorname{Im}(f)) = \dim(E)$

Démontrons les trois premières propriétés :

- Ker $(f) \subset E$ par définition, et $0_E \in \operatorname{Ker}(f)$ car $f(0_E) = 0_F$.

Si $u, u' \in \operatorname{Ker}(f)$ et $\alpha \in \mathbf{R}$, alors $u + u'$ et $\alpha \cdot u \in \operatorname{Ker}(f)$, car

$$f(u + u') = f(u) + f(u') = 0_F + 0_F = 0_F \ ;$$
$$f(\alpha \cdot u) = \alpha \cdot f(u) = \alpha \cdot 0_F = 0_F \ .$$

- Soit f injective, et $u \in \operatorname{Ker}(f)$. $f(u) = f(0_E) = 0_F$, donc $u = 0_E$, donc $\operatorname{Ker}(f) = \{0_E\}$.

Réciproquement, supposons $\text{Ker}(f) = \{0_E\}$. Alors

$$\begin{aligned} f(u) = f\left(u'\right) &\Rightarrow f(u) - f\left(u'\right) = 0 \\ &\Rightarrow f\left(u - u'\right) = 0_F \\ &\Rightarrow u - u' \in \text{Ker}(f) \\ &\Rightarrow u - u' = 0_E \\ &\Rightarrow u = u' \ : \ f \text{ est injective.} \end{aligned}$$

• Par définition, $\text{Im}(f) \subset F$, et $f(0_E) = 0_F$, donc $0_F \in \text{Im}(f)$. Si $v, v' \in \text{Im}(f)$ et $\alpha \in \mathbf{R}$, alors

$$v + v' = f(u) + f\left(u'\right) = f\left(u + u'\right) \in \text{Im}(f)$$

$$\alpha \cdot v = \alpha \cdot f(u) = f(\alpha \cdot u) \in \text{Im}(f)$$

La quatrième propriété est évidente. Pour la cinquième, il suffit d'écrire :

$$v \in \text{Im}(f) \Leftrightarrow v = f(u) = f\left(\sum_{i=1}^{p} \alpha_i \cdot e_i\right) = \sum_{i=1}^{p} \alpha_i \cdot f(e_i)$$

On admet la dernière propriété.

4.2 Applications

Caractérisation des isomorphismes

• Soit $f : E \to E$ un endomorphisme. On a alors les équivalences :

$$\begin{aligned} f \text{ bijective} &\Leftrightarrow f \text{injective} \Leftrightarrow \text{Ker}(f) = \{0_E\} \\ &\Leftrightarrow f \text{surjective} \Leftrightarrow \text{Im}(f) = E \end{aligned}$$

• Soit $f : E \to F$ linéaire. On a les équivalences :

f bijective $\Leftrightarrow f$ injective $(\Leftrightarrow \text{Ker}(f) = \{0_E\})$ et $\dim(E) = \dim(F)$

Démonstration. Il suffit d'appliquer la formule du rang :

$$\dim(\text{Ker}(f)) + \dim(\text{Im}(f)) = \dim(E)$$

Si par exemple f est injective, on a successivement

$$\text{Ker}(f) = \{0_E\}, \dim(\text{Ker}(f)) = 0, \dim(\text{Im}(f)) = \dim(E)$$

Dans le cas d'un endomorphisme, cela implique : $\text{Im}(f) = E$, puis f surjective, puis f bijective. Dans le cas général, si $\dim(E) = \dim(F)$, cela implique $\text{Im}(f) = F$, f surjective, f bijective.

• Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique de $\mathbf{R}^3$ est $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

$$(x, y, z) \in \operatorname{Ker}(f) \Leftrightarrow \begin{cases} x + z = 0 \\ x + y = 0 \\ y + z = 0 \end{cases} \Leftrightarrow \ldots$$
$$\Leftrightarrow \begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases}.$$

$\operatorname{Ker}(f) = \{0\}$, l'**endomorphisme** f est donc bijectif, et on a $\operatorname{Im}(f) = \mathbf{R}^3$.

Attention, on peut avoir $\operatorname{Ker}(f) = \{0_E\}$ sans que f soit bijective. Il faut donc bien préciser que f est un endomorphisme.

Remarques

• $f : E \to F$ linéaire est un isomorphisme ssi l'image d'une base de E est une base de F. En effet, avec $(e_i)_{1 \leqslant i \leqslant n}$ base de E, si f est un isomorphisme, pour tout $v \in F$, il existe un unique $u = \sum_{i=1}^n \alpha_i e_i$ dans E tel que $v = f(u) = \sum_{i=1}^n \alpha_i f(e_i)$, ce qui prouve que $(f(e_i))_{1 \leqslant i \leqslant n}$ est une base de F. Réciproquement, si $(f(e_i))_{1 \leqslant i \leqslant n}$ est une base de F, alors pour tout $v \in F$, il existe $(\alpha_i)_{1 \leqslant i \leqslant n}$ unique dans $\mathbf{R}^n$ tel que $v = \sum_{i=1}^n \alpha_i f(e_i) = f\left(\sum_{i=1}^n \alpha_i e_i\right)$, ce qui prouve que f est un isomorphisme.

• Deux ev E et F sont dits **isomorphes** ssi il existe un isomorphisme $f : E \to F$. Pour que E et F de dimension finie soient isomorphes, il est nécessaire qu'ils soient de même dimension. Cette condition est aussi suffisante : avec $(e_i)_{1 \leqslant i \leqslant n}$, $\left(e'_i\right)_{1 \leqslant i \leqslant n}$ bases respectives de E, F, l'application linéaire f telle que pour tout i, $f(e_i) = e'_i$ est un isomorphisme, d'après la remarque précédente.

Détermination de l'image d'une application linéaire

Avec $f : E \to F$ linéaire, $(e_i)_{1 \leqslant i \leqslant n}$ base de E, on sait que $\operatorname{Im}(f)$ est engendrée par $f(e_1), \ldots, f(e_n)$.

Si la dimension de $\operatorname{Ker}(f)$ est connue, on connaît alors grâce à la formule du rang la dimension de $\operatorname{Im}(f)$, puis une base de $\operatorname{Im}(f)$.

Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique est $A = \begin{pmatrix} 2 & -2 & -2 \\ -1 & 3 & 1 \\ -1 & -1 & 1 \end{pmatrix}$. Pour déterminer $\operatorname{Ker}(f)$, on résout le système homogène $f((x, y, z)) = (0, 0, 0)$, et on trouve
$\operatorname{Ker}(f) = \operatorname{Vect}((1, 0, 1))$, qui est donc de dimension 1. D'après la formule du rang, $\operatorname{Im}(f)$ est de dimension 2. La famille

$$((2, -1, -1), (-2, 3, -1), (-2, 1, 1))$$

est une famille génératrice de $\operatorname{Im}(f)$. On en extrait la famille

$$((2, -1, -1), (-2, 3, -1))$$

(par exemple) libre et de cardinal 2, qui est une donc une base de $\operatorname{Im}(f)$.

On peut aussi déterminer une base de $\operatorname{Im}(f)$ sans passer par le noyau de f : il suffit d'extraire de la famille $(f(e_1), \ldots, f(e_n))$ (qui est une famille génératrice de $\operatorname{Im}(f)$) une sous-famille libre de cardinal maximal. Cette famille est alors une base de $\operatorname{Im}(f)$.

• E est un ev rapporté à une base $\mathcal{B} = (e_1, e_2, e_3)$, a un nombre réel, f_a l'endomorphisme de $\mathbf{R}^3$ tel que :

$$f_a(e_2) = 0 \ ; \ f_a(e_1) = f_a(e_3) = u_1 = a \cdot e_1 + e_2 - a \cdot e_3$$

$\operatorname{Im}(f_a)$ est engendré par la famille $(0, u_1, u_1)$. On en extrait la famille libre de cardinal maximal (u_1), qui est une base de $\operatorname{Im}(f_a)$.

• Soit $\mathcal{B} = (E_1, E_2, E_3, E_4)$ la base canonique de $\mathcal{M}_2(\mathbf{R})$ (voir § 5.1.3), $I = E_1 + E_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $J = E_2 + E_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, f l'application qui à toute matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ associe la matrice

$$f(M) = \frac{a+d}{2} I + \frac{b+c}{2} J.$$

– f est un endomorphisme de $\mathcal{M}_2(\mathbf{R})$ (appliquer la définition, sans oublier de dire que pour tout M de $\mathcal{M}_2(\mathbf{R})$, $f(M)$ appartient à $\mathcal{M}_2(\mathbf{R})$).

– $f(E_1) = f(E_4) = \frac{1}{2} I$, $f(E_2) = f(E_3) = \frac{1}{2} J$. La matrice de f dans la base $\mathcal{B}$ est donc

$$A = \begin{pmatrix} 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}.$$

– Pour déterminer $\text{Im}(f)$, on peut écrire

$$\begin{aligned}\text{Im}(f) &= \text{Vect}(f(E_1), f(E_2), f(E_3), f(E_4)) \\ &= \text{Vect}\left(\frac{1}{2}I, \frac{1}{2}J, \frac{1}{2}J, \frac{1}{2}I\right) = \text{Vect}\left(\frac{1}{2}I, \frac{1}{2}J\right) \\ &= \text{Vect}(I, J)\end{aligned}$$

En effet, $\frac{1}{2}I$ et $\frac{1}{2}J$ ne sont pas colinéaires, ils constituent donc une base de $\text{Im}(f)$ qui est donc de dimension 2. I et J appartiennent à $\text{Im}(f)$ et ne sont pas colinéaires, (I, J) est donc un base de $\text{Im}(f)$.

La dimension de $\text{Im}(f)$ est appelée **rang** de f, noté $\text{rg}(f)$. La formule du rang s'écrit donc :

$$\dim(\text{Ker}(f)) + \text{rg}(f) = \dim(E)$$

Soit f un endomorphisme de $\mathbf{R}^3$. Si f est de rang 0, alors f est l'endomorphisme nul. Si f est de rang 1, les trois vecteurs-colonnes de la matrice A de f dans une base quelconque ont leurs coordonnées proportionnelles. Si f est de rang 2, un des vecteurs-colonnes est colinéaire à une combinaison linéaire des deux autres, qui sont non colinéaires. Si f est de rang 3, alors les trois vecteurs-colonnes de A forment une famille libre ; f est un isomorphisme.

5. Deux applications

5.1 Application aux suites récurrentes linéaires

On va utiliser les outils d'algèbre linéaire rencontrés jusqu'ici (sous-espaces vectoriels, isomorphisme, dimension, bases) pour établir le théorème sur les suites récurrentes linéaires à deux termes (voir § 0.4.4) :

Soit $a, b \in \mathbf{R}$, $(u_n)_{n\in\mathbf{N}}$ telles que $\forall n \in \mathbf{N}, u_{n+2} = au_{n+1} + bu_n$

On considère l'équation caractéristique (1) $r^2 = ar + b$.

• Si l'équation (1) a deux racines r_1, r_2, alors il existe $\alpha, \beta \in \mathbf{N}$ tels que :

$$\forall n \in \mathbf{N}, \quad u_n = \alpha {r_1}^n + \beta {r_2}^n$$

• Si l'équation (1) a une racine double r_1, alors il existe $\alpha, \beta \in \mathbf{N}$ tels que :

$$\forall n \in \mathbf{N}, \quad u_n = \alpha {r_1}^n + \beta n {r_1}^n$$

La démonstration n'est pas détaillée, on en donne juste les grandes lignes :

• L'ensemble $E = \{(u_n)\ ;\ \forall n \in \mathbf{N}, u_{n+2} = au_{n+1} + bu_n\}$ est un espace vectoriel, car c'est un sev de l'ev de référence $\mathbf{R}^{\mathbf{N}}$.

• L'application $\Phi\ :\ E \to \mathbf{R}^2, (u_n) \mapsto (u_0, u_1)$ est un isomorphisme d'espaces vectoriels, car elle est linéaire (revenir à la définition), et bijective (idem).

• D'après la formule du rang, E est donc de dimension 2. On cherche alors deux éléments non colinéaires de E, qui formeront une base de E.

• On cherche d'abord les suites géométriques (r^n) appartenant à E, ce qui a lieu si et seulement si

$$\forall n \in \mathbf{N}, \quad r^{n+2} = ar^{n+1} + br^n \Leftrightarrow r^n\left(r^2 - ar - b\right) = 0$$

• Ainsi apparaît l'équation caractéristique. Si celle-ci a deux solutions r_1 et r_2, les suites $({r_1}^n)$ et $({r_2}^n)$ forment une famille libre, donc une base, de E qui est de dimension 2. Cela veut dire exactement :

$\exists \alpha, \beta \in \mathbf{R}, (u_n) = \alpha \cdot ({r_1}^n) + \beta \cdot ({r_2}^n)$, c'est-à-dire :

$$\forall n \in \mathbf{N}, \quad u_n = \alpha {r_1}^n + \beta {r_2}^n,$$

soit le résultat dans ce cas là.

Enfin, si l'équation caractéristique a une racine double r_1, on vérifie que les suites $({r_1}^n)$ et $(n{r_1}^n)$ appartiennent à E et forment une famille libre, ce qui donne le résultat dans ce cas là.

5.2 Inversibilité du tableau de Pascal

Soit $n \in \mathbf{N}^*$ et φ l'application de $\mathbf{R}_n[X]$ dans $\mathbf{R}[X]$ définie par

$$\varphi(P) = Q \quad \text{tel que} \quad Q(X) = \varphi(P)(X) = P(X+1).$$

• φ est un endomorphisme de $\mathbf{R}_n[X]$. En effet :

– Si P est un polynôme de degré $\leqslant n$, alors $Q(X) = P(X+1)$ est un polynôme de degré $\leqslant n$. φ est donc une application de $\mathbf{R}_n[X]$ dans $\mathbf{R}_n[X]$.

– Si $P_1, P_2 \in \mathbf{R}_n[X]$ et $\alpha \in \mathbf{R}$, alors

$$\begin{aligned}\varphi(P_1 + P_2)(X) &= (P_1 + P_2)(X+1) = P_1(X+1) + P_2(X+1) \\ &= \varphi(P_1)(X) + \varphi(P_2)(X),\end{aligned}$$

donc $\varphi(P_1 + P_2) = \varphi(P_1) + \varphi(P_2)$;

$$\varphi(\alpha P_1)(X) = (\alpha P_1)(X+1) = \alpha P_1(X+1) = \alpha\varphi(P_1)(X)$$

donc φ est linéaire.

• φ est un automorphisme de $\mathbf{R}_n[X]$, et $\varphi^{-1}(Q)(X) = Q(X-1)$ pour tout $Q \in \mathbf{R}_n[X]$. En effet :

$$Q = \varphi(P) \Leftrightarrow Q(X) = P(X+1) \Leftrightarrow P(X) = Q(X-1)$$

• soit A la matrice de φ dans la base canonique de $\mathbf{R}_n[X]$. Les colonnes de A sont les coordonnées de $\varphi(1), \varphi(X), \ldots, \varphi(X^n)$ dans la base $(1, X, \ldots, X^n)$. Attention, il s'agit des (fonctions) polynômes $1, X, \ldots,$ et non des nombres réels. On a

$$\varphi(1) = 1 \ ; \ \varphi(X) = X+1 \ ; \ \ldots \ ; \ \varphi(X^n) = (X+1)^n$$

En effet, si, pour tout $X \in \mathbf{R}$, $P(X) = 1$, alors, pour tout $X \in \mathbf{R}$,

$$Q(X) = \varphi(P)(X) = P(X+1) = 1, \text{ etc.}$$

On utilise alors la formule du binôme :

$$(X+1)^k = \sum_{i=0}^{k} \binom{k}{i} X^i$$

La matrice A est donc la matrice

$$A = \begin{pmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \cdots & \cdots & \binom{n}{0} \\ 0 & \binom{1}{1} & \binom{2}{1} & \cdots & \cdots & \binom{n}{1} \\ \vdots & 0 & \binom{2}{2} & \cdots & \cdots & \binom{n}{2} \\ & \vdots & 0 & \ddots & & \vdots \\ & & \vdots & \ddots & \ddots & \\ 0 & 0 & 0 & & 0 & \binom{n}{n} \end{pmatrix}$$

• Puisque φ est un automorphisme, A est inversible, et A^{-1} est la matrice de φ^{-1} dans la même base $(1, X, \ldots, X^n)$. On a

$$\varphi^{-1}\left(X^k\right) = (X-1)^k = \sum_{i=0}^{k} \binom{k}{i} X^i (-1)^{k-i}$$

et par conséquent :

$$A^{-1} = \begin{pmatrix} \binom{0}{0} & -\binom{1}{0} & \binom{2}{0} & \cdots & \cdots & (-1)^n \binom{n}{0} \\ 0 & \binom{1}{1} & -\binom{2}{1} & \cdots & \cdots & (-1)^{n-1} \binom{n}{1} \\ \vdots & 0 & \binom{2}{2} & \cdots & \cdots & (-1)^{n-2} \binom{n}{2} \\ & \vdots & 0 & \ddots & & \vdots \\ & & \vdots & \ddots & \ddots & \\ 0 & 0 & 0 & & 0 & \binom{n}{n} \end{pmatrix}$$

• A est la transposée du tableau de Pascal considéré comme une matrice. On a donc bien obtenu l'inverse du tableau de Pascal d'ordre n. Par exemple, avec $n = 4$:

$${}^tA = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix} \; ; \; \left({}^tA\right)^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix}$$

On vérifie que ça marche (le produit des deux matrices est égal à I_4).



Diagonalisation 6

Les matrices carrées avec lesquelles les calculs sont les plus simples sont les matrices diagonales. On a vu en effet que, si $p \in \mathbf{N}^*$ et $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$, alors $D^p = \begin{pmatrix} \lambda_1^p & 0 & 0 \\ 0 & \lambda_2^p & 0 \\ 0 & 0 & \lambda_3^p \end{pmatrix}$, ceci restant valable avec $p = -1$ (c'est-à-dire avec D est inversible) ssi tous les λ_i sont non nuls.

Le problème de la diagonalisation peut se poser de la manière suivante : soit f un endomorphisme de l'ev E de dimension n, $\mathcal{B}$ une base de E, $A = \text{Mat}\,(f, \mathcal{B})$ la matrice de f dans la base $\mathcal{B}$. Trouver, si elle existe, une base $\mathcal{B}'$ de E dans laquelle la matrice de f soit diagonale.

Le § 2 est l'objet de cette étude. Auparavant, on doit savoir « passer » d'une base à une autre, c'est l'objet du § 1. Les résultats seront donnés dans le cadre abstrait de l'ev E de dimension n, muni d'une base $\mathcal{B}$. Dans les exemples des deux premiers paragraphes, on aura toujours $E = \mathbf{R}^3$, $\mathcal{B}$ la base canonique de $\mathbf{R}^3$.

1. Théorie du changement de base

Définition. Soit $\mathcal{B}$ une base de l'ev E de cardinal n, $\mathcal{F}$ une famille de n vecteurs de E. La **matrice de passage** de la base $\mathcal{B}$ à la famille $\mathcal{F}$ est la matrice P dont les colonnes successives sont les coordonnées des vecteurs $u_1, \ldots, u_n$ dans la base $\mathcal{B}$.

Propriétés

Proposition 1. La matrice P est inversible ssi $\mathcal{F}$ est une base de E. La matrice P^{-1} est alors la matrice de passage de la base $\mathcal{F}$ à la base $\mathcal{B}$.

Exemple. Soit $\mathcal{F} = (u_1, u_2, u_3)$, avec

$$u_1 = (1, 0, 0)\,, \quad u_2 = (2, 2, 0)\,, \quad u_3 = (3, 3, 3)$$

La matrice de passage de $\mathcal{B}$ (base canonique de $\mathbf{R}^3$) à $\mathcal{F}$ est $P = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$. P est une matrice triangulaire sans zéros sur la diagonale, elle est donc inversible, et $\mathcal{F}$ est une base de $\mathbf{R}^3$. En calculant P^{-1}, par exemple grâce à la méthode du pivot, on trouve : $P^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & 1/3 \end{pmatrix}$. P^{-1} est la matrice de passage de la base $\mathcal{F}$ à la base $\mathcal{B}$, les vecteurs colonnes de P^{-1} fournissent donc :

$$e_1 = u_1 \;\; ; \quad e_2 = -u_1 + \frac{1}{2}u_2 \;\; ; \quad e_3 = -\frac{1}{2}u_2 + \frac{1}{3}u_3$$

(On aurait pu procéder dans l'ordre inverse, en exprimant les e_i en fonction des u_i, et en déduire alors P^{-1}.)

- Ne pas confondre la matrice P et la famille $\mathcal{F}$. P est inversible ssi la famille $\mathcal{F}$ est une base (ou une famille libre), mais on ne parle pas de matrice libre ni de famille inversible.
- La proposition 1 donne une nouvelle manière de déterminer si une famille $\mathcal{F}$ est une base (ssi la matrice de passage de $\mathcal{B}$ à $\mathcal{F}$ est inversible).

Proposition 2. (Effet d'un changement de base sur les coordonnées d'un vecteur). Soit $\mathcal{B}, \mathcal{B}'$ deux bases de E, P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, u un vecteur de E, $X = \mathrm{Mat}\,(u, \mathcal{B})$, $X' = \mathrm{Mat}\left(u, \mathcal{B}'\right)$ les matrices colonnes des coordonnées de u dans les bases $\mathcal{B}, \mathcal{B}'$.
Alors

$$X = PX'$$

La matrice de passage est construite en écrivant en colonne les coordonnées des vecteurs de la nouvelle base par rapport à l'ancienne, mais elle donne les anciennes coordonnées des vecteurs par rapport aux nouvelles.

Proposition 3 (Effet d'un changement de base sur la matrice d'un endomorphisme). Soit

- $\mathcal{B}, \mathcal{B}'$ deux bases de E ;
- P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}''$;
- f un endomorphisme de E ;
- $A = \text{Mat}(f, \mathcal{B})$, $A' = \text{Mat}\left(f, \mathcal{B}'\right)$ les matrices de f dans les bases $\mathcal{B}, \mathcal{B}'$.

Alors
$$A' = P^{-1}AP$$

De façon équivalente, on a $A = PA'P^{-1}$, voir la première remarque ci-dessous. C'est d'ailleurs sous cette forme que vous utiliserez le plus souvent la théorie du changement de base, la matrice A' étant plus simple à manipuler que la matrice A.

Par exemple, avec
$$n \in \mathbf{N}^*, A^n = \left(PA'P^{-1}\right) \dots \left(PAP^{-1}\right) = PA'^n P^{-1}.$$

Si A' est diagonale, alors on calcule facilement A'^n, puis A^n.

Matrices semblables

Deux matrices A, A' appartenant à $\mathcal{M}_n(\mathbf{R})$ sont dites **semblables** ssi il existe une matrice inversible P appartenant à $\mathcal{M}_n(\mathbf{R})$ telles que
$$A' = P^{-1}AP$$

Deux matrices sont semblables ssi elles représentent dans deux bases le même endomorphisme de E ev de dimension n.

Remarques

- A est semblable à A, car $A = I_n^{-1}AI_n$.

Si A est semblable à A', alors A' est semblable à A :
$$A' = P^{-1}AP \Rightarrow PA'P^{-1} = P\left(P^{-1}AP\right)P^{-1} = \left(PP^{-1}\right)A\left(PP^{-1}\right) = A$$

Si A est semblable à A' et A' est semblable à A'', alors A est semblable à A'' :
$$A = PA'P^{-1}, A' = QA''Q^{-1} \Rightarrow A = PQA''Q^{-1}P^{-1} = (PQ)\,A''(PQ)^{-1}$$

On dit que la relation « A est semblable à A' » est une relation d'équivalence.

• P étant une matrice inversible fixée, on vérifie sans peine que l'application ϕ_P de $\mathcal{M}_n(\mathbf{R})$ dans $\mathcal{M}_n(\mathbf{R})$ définie par $\phi_P(M) = P^{-1}MP$ est un automorphisme, dont la réciproque est $\phi_P^{-1}(N) = PNP^{-1}$. Il vérifie de plus : $\phi_P(MM') = \phi_P(M)\,\phi_P(M')$ pour tout M, M'.

2. Diagonalisation

2.1 Valeurs propres, vecteurs propres

Soit $f \in \mathcal{L}(E)$, $\mathcal{B}$ une base de E, A la matrice de f dans la base $\mathcal{B}$. Supposons qu'il existe une base $\mathcal{C} = (u_1, \ldots, u_n)$ de E dans laquelle la matrice D de f soit diagonale :

$$D = \mathrm{Mat}(f, \mathcal{C}) = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

On a alors, pour tout $i \in \{1, \ldots, n\}$: $f(u_i) = \lambda_i \cdot u_i$

Cela conduit naturellement aux définitions suivantes.

Définitions

Soit $f \in \mathcal{L}(E)$, $u \in E$, $u \neq 0_E$ et $\lambda \in \mathbf{R}$ tels que $f(u) = \lambda \cdot u$.

On dit alors que λ est une **valeur propre** de f, et que u est un **vecteur propre** de f pour la valeur propre λ (ou : associé à la valeur propre λ).

On précise $u \neq 0_E$, sinon tout λ réel serait valeur propre :

$$\forall \lambda \in R,\ f(0_E) = 0_E = \lambda \cdot 0_E$$

Par contre, 0 peut être valeur propre. C'est le cas ssi f n'est pas injectif.

Soit A une matrice de $\mathcal{M}_n(\mathbf{R})$, X une matrice colonne de $\mathcal{M}_{n,1}(\mathbf{R})$, non nulle, et $\lambda \in \mathbf{R}$, tels que $AX = \lambda X$.

On dit alors que λ est une **valeur propre** de A, et que X est un **vecteur propre** de A pour la valeur propre λ (ou : associé à la valeur propre λ).

On dit aussi : X est un vecteur-colonne propre, ou une matrice colonne propre.

Proposition et définition

> Soit λ une valeur propre de f.
> Alors l'ensemble
> $$F_\lambda = \{u \in E \mid f(u) = \lambda \cdot u\}$$
> est un sev de E, non réduit au vecteur 0_E, appelé le **sous-espace propre** de f associé à la valeur propre λ.

En effet :

$$\begin{aligned} u \in F_\lambda \Leftrightarrow f(u) - \lambda u = 0_E &\Leftrightarrow (f - \lambda \cdot \mathrm{Id}_E)(u) = 0_E \\ &\Leftrightarrow u \in \mathrm{Ker}(f - \lambda \cdot \mathrm{Id}_E) \end{aligned}$$

$f - \lambda \cdot \mathrm{Id}_E$ est un endomorphisme de E, et on sait que le noyau d'une application linéaire est un sev. D'autre part F_λ n'est pas réduit à 0_E car on a supposé que λ était valeur propre de f.

On parlera de même du sous-espace propre de la matrice A associé à la valeur propre λ. Ce sous-espace propre est un sev de $\mathcal{M}_{n,1}(\mathbf{R})$.

On abrégera valeur propre, vecteur propre, sous-espace propre en : vap, vep, sep.

Recherche des valeurs propres et des vecteurs propres

De l'étude du chapitre 5, on déduit et on retient :

> Soit $f \in \mathcal{L}(E)$, $\mathcal{B}$ une base de E, $A = \mathrm{Mat}(f, \mathcal{B})$.
> $$\begin{aligned} \lambda \text{ valeur propre de } f &\Leftrightarrow \lambda \text{ valeur propre de } A \\ &\Leftrightarrow \mathrm{Ker}(f - \lambda \cdot \mathrm{Id}_E) \neq \{0_E\} \\ &\Leftrightarrow f - \lambda \cdot \mathrm{Id}_E \text{ non injective} \\ &\Leftrightarrow f - \lambda \cdot \mathrm{Id}_E \text{ non bijective} \\ &\Leftrightarrow A - \lambda \cdot I_n \text{ non inversible} \\ &\Leftrightarrow \exists X \in \mathcal{M}_n(\mathbf{R}),\ X \neq 0,\ (A - \lambda \mathrm{I}_n)\, X = 0 \end{aligned}$$

Pratiquement, on utilise la dernière équivalence, qui conduit à résoudre un système homogène, en discutant suivant la valeur de λ. On a traité en détail un exemple, § 4.1.3, on en traite ici un autre, de façon plus synthétique.

Déterminons les valeurs propres et vecteurs propres de $f \in \mathcal{L}\left(\mathbf{R}^3\right)$ dont la matrice dans la base canonique de $\mathbf{R}^3$ est :

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

On cherche $\lambda \in \mathbf{R}$ et $X \neq 0$ tels que $(A - \lambda \cdot \mathrm{I}_3)\, X = 0$. Par la méthode du pivot :

$$\begin{pmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{pmatrix} L_1 \leftrightarrow L_3$$

$$\begin{pmatrix} 1 & 0 & 1-\lambda \\ 0 & 2-\lambda & 0 \\ 1-\lambda & 0 & 1 \end{pmatrix} L_3 \leftarrow L_3 - (1-\lambda)\, L_1$$

$$\begin{pmatrix} 1 & 0 & 1-\lambda \\ 0 & 2-\lambda & 0 \\ 0 & 0 & \lambda\,(2-\lambda) \end{pmatrix}$$

Les valeurs propres de f sont donc 0 et 2.

Pour $\lambda = 0$:

$$\begin{cases} x + z = 0 \\ 2y = 0 \\ 0 = 0 \end{cases} \Leftrightarrow \begin{cases} x = -z \\ y = 0 \end{cases}$$

solutions $(x, y, z) = (-z, 0, z) = z\,(-1, 0, 1)$

Le sous-espace propre associé est $F_0 = \mathrm{Vect}\,((-1, 0, 1))$.

Pour $\lambda = 2$:

$$\begin{cases} x - z = 0 \\ 0 = 0 \\ 0 = 0 \end{cases} \Leftrightarrow x = z$$

solutions $(x, y, z) = (z, y, z) = z\,(1, 0, 1) + y\,(0, 1, 0)$

Le sous-espace propre associé est $F_2 = \mathrm{Vect}\,((1, 0, 1), (0, 1, 0))$.

• On peut aussi dire : les valeurs propres de A sont 0 et 2. Le sous-espace propre de A pour la valeur propre 0 est $\mathrm{Vect}\left(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right)$ $\left(\text{ou } \mathrm{Vect}\,((-1, 0, 1))\right)$, etc.

• N'écrivez pas sur votre copie les calculs annexes (mais faites-les !). Par exemple l'opération élémentaire $L_3 \leftarrow L_3 - (1-\lambda)\, L_1$ donne

lieu aux calculs

$$1 - (1-\lambda)^2 = (1 - 1 + \lambda)(1 + 1 - \lambda) = \lambda(2-\lambda)$$

Par contre vous mentionnerez bien les opérations élémentaires sur les lignes, en proscrivant absolument toute opération du type $L_i \leftarrow \alpha L_i + \beta L_j$, avec α susceptible d'être nul.

• Les résultats obtenus sont susceptibles d'être vérifiés, et il est conseillé de le faire ! Dans l'exemple traité, la vérification est :

$$A\begin{pmatrix}-1\\1\\0\end{pmatrix} = 0\begin{pmatrix}-1\\1\\0\end{pmatrix} \;;\quad A\begin{pmatrix}1\\0\\1\end{pmatrix} = 2\begin{pmatrix}1\\0\\1\end{pmatrix} \;;\quad A\begin{pmatrix}0\\1\\0\end{pmatrix} = 2\begin{pmatrix}0\\1\\0\end{pmatrix}$$

Et il vaut mieux prendre le temps de corriger une erreur que de persister dans celle-ci, surtout si le traitement des questions suivantes dépend de la bonne réponse ...

• Le savoir technique mis en œuvre ici ne doit pas vous faire oublier la définition d'un vecteur propre et d'une valeur propre ($f(u) = \lambda \cdot u$ en abrégé), qu'il est essentiel de bien comprendre. Par exemple, si on demande de vérifier que $u = (1, 0, 2)$ est un vecteur propre pour l'endomorphisme f de $\mathbf{R}^3$ de matrice $A = \begin{pmatrix}-19 & 6 & 9\\-30 & 11 & 15\\-22 & 6 & 10\end{pmatrix}$ dans la base canonique, il serait très maladroit de suivre la démarche précédente, alors qu'il suffit d'effectuer

$$\begin{pmatrix}-19 & 6 & 9\\-30 & 11 & 15\\-22 & 6 & 10\end{pmatrix}\begin{pmatrix}1\\0\\2\end{pmatrix} = \begin{pmatrix}-1\\0\\-2\end{pmatrix} = (-1)\begin{pmatrix}1\\0\\2\end{pmatrix}$$

pour pouvoir dire : $f(u) = -u$, donc u est un vecteur propre de f pour la valeur propre 1.

2.2 Endomorphismes et matrices diagonalisables

Définitions

L'endomorphisme f de $\mathcal{L}(E)$ est dit **diagonalisable** ssi il existe une base de E dans laquelle la matrice de f est diagonale, ou de façon équivalente :

Un endomorphisme f de E est diagonalisable ssi il existe une base de E formée de vecteurs propres de f.

Soit $A = \mathrm{Mat}(f, \mathcal{B})$. On dit que la matrice A est diagonalisable ssi f est diagonalisable. D'après la théorie du changement de base, on a de façon équivalente :

La matrice A de $\mathcal{M}_n(\mathbf{R})$ est diagonalisable ssi il existe une matrice diagonale D et une matrice inversible P de $\mathcal{M}_n(\mathbf{R})$ telles que

$$A = PDP^{-1}$$

Vous ferez un effort particulier pour comprendre et mémoriser ces définitions encadrées.

Critères pour la diagonalisation

Théorème. Soient $u_1, \ldots, u_p$ des vecteurs propres associés à des valeurs propres distinctes $\lambda_1, \ldots, \lambda_p$ de l'endomorphisme f. Alors la famille $(u_1, \ldots, u_p)$ est libre.

Démontrons ce théorème dans le cas particulier où $p = 2$. Soient α_1, α_2 des réels tels que $\alpha_1 \cdot u_1 + \alpha_2 \cdot u_2 = 0_E$. En appliquant f on trouve $\alpha_1\lambda_1 \cdot u_1 + \alpha_2\lambda_2 \cdot u_2 = 0_E$. Mais $\alpha_2 \cdot u_2 = -\alpha_1 \cdot u_1$, donc

$$\alpha_1\lambda_1 \cdot u_1 - \alpha_1\lambda_2 \cdot u_1 = 0_E, \alpha_1(\lambda_1 - \lambda_2) \cdot u_1 = 0_E.$$

u_1 est non nul, λ_1 est différent de λ_2, c'est donc α_1 qui est nul, puis α_2 aussi, car u_2 est non nul. La famille (u_1, u_2) est donc libre.

On démontre ce théorème dans le cas général par récurrence.

On déduit immédiatement de ce théorème une condition suffisante de diagonalisation :

Soit f un endomorphisme de E, ev de dimension n. Si f admet n valeurs propres distinctes, alors f est diagonalisable.

En effet, les vecteurs $u_1, \ldots, u_n$ associés aux n valeurs propres distinctes forment une famille libre. Comme elle est de cardinal n, et que E est de dimension n, c'est une base de E formée de vecteurs propres de f.

Attention, il s'agit d'une condition suffisante pour que f soit diagonalisable, mais pas nécessaire : il existe des endomorphismes de $\mathbf{R}^3$ diagonalisables avec 2 valeurs propres distinctes, ou une seule.

On évitera donc d'écrire « $f \in \mathcal{L}(\mathbf{R}^3)$ n'admet pas trois valeurs propres distinctes, f n'est donc pas diagonalisable. »

Voici une condition nécessaire et suffisante de diagonalisation :

Soit f un endomorphisme de E, ev de dimension n. f est diagonalisable ssi la somme des dimensions des sous-espaces propres de f est égale à n.

Propriété admise.

On a les théorèmes analogues pour la matrice A de $\mathcal{M}_n(\mathbf{R})$:

Les vecteurs colonnes propres de A associés à des valeur propres distinctes de A forment une famille libre de $\mathcal{M}_{n,1}(\mathbf{R})$.

Si A admet n valeur propres distinctes, alors A est diagonalisable.

A est diagonalisable ssi la somme des dimensions des sous-espaces propres de A est égale à n.

• Soit $f \in \mathcal{L}\left(\mathbf{R}^3\right)$ de matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ dans la base canonique de $\mathbf{R}^3$. f admet 3 valeurs propres distinctes, les réels $1, 2, 3$. (En effet, λ est valeur propre de f ssi $A - \lambda I_3$ n'est pas inversible.) Donc f est diagonalisable. En résolvant les systèmes triangulaires $(A - \lambda I_3) X = 0$ avec $\lambda = 1, 2, 3$, on touve respectivement les sep $\text{Vect}(u_1)$, $\text{Vect}(u_2)$, $\text{Vect}(u_3)$, avec

$$u_1 = (1, 0, 0), u_2 = (1, 1, 0), u_3 = (3, 4, 2)$$

$\mathcal{C} = (u_1, u_2, u_3)$ est une base de $\mathbf{R}^3$ formée de vecteurs propres de f.

La matrice de f dans cette base est $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

D'après la théorie du changment de base, la matrice $P = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}$ est inversible, et $A = PDP^{-1}$.

• Avec $f \in \mathcal{L}\left(\mathbf{R}^3\right)$ de matrice $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ dans la base canonique de $\mathbf{R}^3$, on a trouvé précédemment :

0 vap de f, sep associé $F_0 = \text{Vect}((-1, 0, 1))$.

2 vap de f, sep associé $F_2 = \text{Vect}((1, 0, 1), (0, 1, 0))$

La somme des dimensions des sous-espaces propres de f est égale à $1 + 2 = 3$, donc f est diagonalisable.

$\mathcal{C} = ((-1, 0, 1), (1, 0, 1), (0, 1, 0))$ est une base de $\mathbf{R}^3$ formée de vecteurs propres de f. La matrice de f dans cette base est $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

D'après la théorie du changment de base, la matrice $P = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ est inversible, et $A = PDP^{-1}$.

Notez dans ces deux exemples comment on évoque la théorie du changement de base, ce qui permet de montrer rapidement que la matrice P est effectivement inversible : P est inversible car c'est la matrice de passage de la base canonique $\mathcal{B}$ à la base $\mathcal{C}$, et $\mathcal{C}$ est une base de $\mathbf{R}^3$ car f (ou A) est diagonalisable.

• Soit $f \in \mathcal{L}(\mathbf{R}^3)$ diagonalisable avec une seule valeur propre λ. Alors la dimension du sep associé est égale à 3, le sep associé est $\mathbf{R}^3$ lui-même. Donc pour tout $u \in \mathbf{R}^3, f(u) = \lambda \cdot u$. La matrice de f dans toute base de $\mathbf{R}^3$ est λI_3. En dehors de ce cas, un endomorphisme de $\mathbf{R}^3$ possèdant une seule valeur propre n'est pas diagonalisable.

Soit $A \in \mathcal{M}_3(\mathbf{R})$ diagonalisable avec une seule valeur propre λ. Alors il existe P inversible telle que $A = P(\lambda I_3) P^{-1} = \lambda I_3$. En dehors de ce cas, une matrice avec une seule valeur propre n'est pas diagonalisable.

• Avec $f \in \mathcal{L}(\mathbf{R}^3)$ de matrice $A = \begin{pmatrix} -2 & 5 & 7 \\ -1 & 6 & 9 \\ 0 & -2 & -3 \end{pmatrix}$ dans la base canonique de $\mathbf{R}^3$, on trouve

-1 vap de f, sep associé $\text{Vect}((2, -1, 1))$.

1 vap de f, sep associé $\text{Vect}((1, 2, -1))$.

La somme des dimensions des sous-espaces propres est égale à 2, donc $f \in \mathcal{L}(\mathbf{R}^3)$ n'est pas diagonalisable. (Le fait que f ait deux valeurs propres ne suffit pas pour conclure.)

Propriétés diverses

Soit f un endomorphisme de l'ev E de dimension n. Alors f admet au plus n valeurs propres distinctes.

$A \in \mathcal{M}_n(\mathbf{R})$ admet au plus n valeurs propres distinctes.

En effet, une famille libre dans un ev de dimension n est de cardinal au plus n.

Si, dans la recherche des valeurs propres de $f \in \mathcal{L}\left(\mathbf{R}^3\right)$, vous trouvez 4 valeurs propres, il y a une erreur, à ne pas laisser passer !

On peut généraliser ce résultat en disant que la somme des dimensions des sous-espaces propres d'un endomorphisme de E de dimension n est au plus égale à n. Ainsi, si l'énoncé met en évidence, pour $f \in \mathcal{L}\left(\mathbf{R}^3\right)$, deux valeurs propres, les sous-espaces propres associés étant de dimension 1 et 2, il est inutile de chercher d'autres valeurs propres, et on peut conclure que f est diagonalisable.

Les valeurs propres d'une matrice triangulaire sont les éléments de sa diagonale.

En effet, λ est valeur propre de A ssi $A - \lambda I_n$ n'est pas inversible. Si A est triangulaire, $A - \lambda I_n$ l'est aussi, et une matrice triangulaire est inversible ssi il n'y a pas de zéros sur sa diagonale.

L'endomorphisme f est bijectif ssi 0 n'est pas valeur propre de f.

La matrice A est inversible ssi 0 n'est pas valeur propre de A.

En effet, λ est valeur propre de f ssi $f - \lambda \cdot \mathrm{Id}_E$ n'est pas bijectif.

• Propriété très importante, qui correspond à une question souvent posée. Si on connaît l'ensemble des valeurs propres de f, alors il est inutile d'utiliser l'algorithme du pivot, ou toute autre méthode, pour dire si f est bijectif.

• Ne confondez pas matrice diagonalisable et matrice inversible ! Il n'y a aucune relation d'implication entre ces deux propriétés.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ est diagonalisable et inversible, $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ diagonalisable et non inversible, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ non diagonalisable (sinon, avec 1 pour seule valeur propre, la matrice serait égale à $PI_3P^{-1} = I$) et inversible (0 n'est pas valeur propre), $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ non diagonalisable et non inversible.

Toute matrice symétrique réelle est diagonalisable.

Propriété admise, et qui ne peut donner lieu qu'à une question de cours (« montrer sans calculs que A est diagonalisable »).

On dit que le polynôme $P(X)$ est un **polynôme annulateur** de la matrice carrée A ssi $P(A) = 0$.

Soit $P(X)$ un polynôme annulateur de A. Alors toute valeur propre de A est racine de ce polynôme.

• Soit A une matrice de $\mathcal{M}_3(\mathbf{R})$, différente de I_3, telle que

$$A^3 - A^2 + A - I_3 = O_3.$$

A admet pour polynôme annulateur le polynôme

$$P(X) = X^3 - X^2 + X - 1$$

(En effet, $1 = X^0$. En remplaçant X par A, on obtient $A^0 = I_3$.)

Toute valeur propre de A est racine de $P(X)$.

Or $P(X) = (X-1)\left(X^2+1\right)$ admet pour seule racine le nombre réel 1. On ne sait pas si 1 est effectivement valeur propre de A, mais on peut dire que A n'est pas diagonalisable : sinon, sa seule valeur propre serait 1, on aurait donc $A = PI_3P^{-1} = I$, ce qu'on a supposé être faux.

Remarquons d'autre part que A est inversible et que le polynôme annulateur permet d'exprimer A^{-1} en fonction de A :

$A^3 - A^2 + A - I_3 = O_3$, donc

$A^3 - A^2 + A = I_3$, $A\left(A^2 - A + I_3\right) = I_3$, donc A est inversible, et

$A^{-1} = A^2 - A + I_3$.

• Soit $A = \begin{pmatrix} 1 & 1 & -1 & -3 \\ 1 & 1 & 1 & -2 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & 0 & -2 \end{pmatrix}$. $A^2 = -I_4$, donc $A^2 + I_4 = O_4$.

A admet le polynôme annulateur $X^2 + 1$, qui n'a pas de racine réelle. Donc A n'admet pas de valeur propre réelle. A n'est donc pas diagonalisable, et elle est inversible (sinon, elle aurait 0 pour valeur propre). On peut préciser ce dernier résultat en écrivant :

$A^2 = -I_4$, donc $A(-A) = I_4$, donc A est inversible, et $A^{-1} = -A$.

• Soit $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$. $A^2 \neq O_3, A^3 = O_3$. A admet pour polynôme annulateur $P(X) = X^3$, dont la seule racine est 0. Donc si λ est valeur propre de A, alors $\lambda = 0$. Attention, cela ne prouve pas que 0 est effectivement valeur propre de A. Mais A n'est pas inversible (sinon on aurait $A^{-1}A^3 = O_3 A^3 = O_3, \mathrm{A}^2 = O_3$, or $A^2 \neq O_3$), donc 0 est valeur propre de A, et c'est la seule, donc A n'est pas diagonalisable (sinon aurait $A = PO_3P^{-1} = O_3$).

On voit sur ces trois exemples l'utilisation qui est faite d'un polynôme annulateur de A pour déterminer si A est inversible, et calculer éventuellement son inverse. On peut aussi utiliser un polynôme annulateur de A pour calculer A^n, c'est ce qu'on a fait dans l'exemple 2 du § 4.2.4, où le polynôme annulateur était

$$P(X) = X^3 - X^2 - 2X.$$

On rappelle que si $A = \mathrm{Mat}(f, \mathcal{B})$, alors $A^n = \mathrm{Mat}(f^n, \mathcal{B})$, avec $f^0 = \mathrm{Id}_E$, et $f^n = f \circ \cdots \circ f$ (n termes). Un polynôme annulateur de A est donc aussi un polynôme annulateur de f, et on peut dire :

Si l'endomorphisme f admet un polynôme annulateur P, alors toute valeur propre de f est racine de ce polynôme.

Démontrons cette propriété. Soient $P(X) = \sum_{k=0}^{p} a_k X^k$ un polynôme annulateur de f (donc $\sum_{k=0}^{p} a_k f^k(v) = 0_E$ pour tout v), $u \neq 0_E$ et $\lambda \in \mathbf{R}$ tels que $f(u) = \lambda \cdot u$. Alors

$$f^2(u) = f(f(u)) = f(\lambda u) = \lambda f(u) = \lambda\lambda u = \lambda^2 u,$$

puis, par récurrence : $f^k(u) = \lambda^k u$, puis $\sum_{k=0}^{p} a_k \lambda^k u = 0_E$, puis $\sum_{k=0}^{p} a_k \lambda^k = 0$ car $u \neq 0_E$, $P(\lambda) = 0$.

3. Autres réductions — Applications

3.1 Autres réductions

Si l'endomorphisme f (resp. la matrice carrée A) n'est pas diagonalisable, on peut chercher une base de E (resp. une matrice P inversible) telle que la matrice de f dans cette base (resp. la matrice $A' = PAP^{-1}$) soit « simple », en général triangulaire. Aucune connaissance spécifique n'est exigible, on donne quelques exemples.

• On a vu que l'endomorphisme f de $\mathbf{R}^3$ de matrice $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$ dans la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ n'était pas diagonalisable (voir deuxième exemple du § 6.2.2).

Soit $\varepsilon_1 = e_1, \varepsilon_2 = f(\varepsilon_1), \varepsilon_3 = f(\varepsilon_2)$, et $\mathcal{C} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$.

$\varepsilon_1 = (1, 0, 0), \varepsilon_2 = (1, 1, 0), \varepsilon_3 = (1, 1, 1)$. La matrice de passage de la base $\mathcal{B}$ à la famille $\mathcal{C}$ est $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Elle est triangulaire sans zéros sur la diagonale, donc inversible, donc $\mathcal{C}$ est une base de $\mathbf{R}^3$. La matrice de f dans la base $\mathcal{C}$ est $N = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, car $f(\varepsilon_1) = \varepsilon_2, f(\varepsilon_2) = \varepsilon_3, f(\varepsilon_3) = 0$. D'après la théorie du changement de base, on a $A = PNP^{-1}$. N n'est pas diagonale, mais triangulaire inférieure, et les propriétés de f se voient clairement sur la matrice N :

Sa seule valeur propre est 0, f n'est donc pas bijective. Et puisque $f(\varepsilon_1) = \varepsilon_2, f(\varepsilon_2) = \varepsilon_3, f(\varepsilon_3) = 0$, on a $N^3 = O_3, f^3 = 0$.

• Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice relativement à la base canonique $\mathcal{B}$ de $\mathbf{R}^3$ est $A = \begin{pmatrix} 3 & -1 & 0 \\ 1 & 6 & 1 \\ -3 & -8 & 0 \end{pmatrix}$. La matrice $N = A - 3I_n$ vérifie $N^2 \neq O_3, N^3 = O_3$. On en déduit comme précédemment que l'endomorphisme $h = f - 3\,\mathrm{Id}_{\mathbf{R}^3}$ admet 0 pour seule valeur propre, puis que f admet 3 pour seule valeur propre :

si $u \neq 0, f(u) = \lambda \cdot u \Leftrightarrow h(u) = (\lambda - 3) \cdot u \Leftrightarrow \lambda - 3 = 0$.

f n'est pas diagonalisable, sinon, avec 3 pour seule valeur propre, on aurait $f(u) = 3 \cdot u$ pour tout $u \in \mathbf{R}^3$.

Mais avec $u_1 = (1, -1, 1), u_2 = h(u_1), u_3 = h(u_2)$,

on vérifie que $\mathcal{B}' = (u_1, u_2, u_3)$ est une base de $\mathbf{R}^3$, puis que la matrice de h dans cette base est $N' = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. L'ensemble des résultats obtenus conduit alors à $A = PA'P^{-1}$, avec

$$A' = 3I_3 + N' = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 2 & -1 \end{pmatrix}.$$

3.2 Applications

Puissances d'une matrice, utilisation en probabilités

Supposons la matrice carrée A diagonalisable : $A = PDP^{-1}$, avec P inversible et D diagonalisable. Alors, pour tout $p \in \mathbf{N}^*$:

$$\begin{aligned} A^p &= \left(PDP^{-1}\right)\ldots\left(PDP^{-1}\right) \text{ (p facteurs)} \\ &= PD^pP^{-1} \end{aligned}$$

(ce qu'on peut aussi établir par récurrence). Le calcul de A^p peut alors être effectué : on calcule D^p, puis P^{-1}, puis PD^pP^{-1}.

> Pour $p \in \left]0, 1/2\right[$ on considère
>
> $$A = \begin{pmatrix} 0 & p & 1-2p & p \\ p & 0 & p & 1-2p \\ 1-2p & p & 0 & p \\ p & 1-2p & p & 0 \end{pmatrix},$$
>
> et
>
> $$P = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$
>
> On a $P^2 = 4I_4,\ P\left(\frac{1}{4}P\right) = I_4$, donc P est inversible et $P^{-1} = \frac{1}{4}P$.
>
> On vérifie que les 4 vecteurs colonnes u_1, u_2, u_3, u_4 de la matrice A sont vecteurs propres de A pour les valeurs propres respectives $1, 1-4p, 2p-1, 2p-1$. Ils constituent donc une base de $\mathcal{M}_{1,4}(\mathbf{R})$ formée de vecteurs propres de A. A est donc diagonalisable, et
>
> $$A = PDP^{-1}, \quad \text{avec} \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-4p & 0 & 0 \\ 0 & 0 & 2p-1 & 0 \\ 0 & 0 & 0 & 2p-1 \end{pmatrix}.$$
>
> Donc $A = \frac{1}{4}PDP, A^n = \left(\frac{1}{4}\right)^n PD^nP$
>
> Le calcul de A^n n'est plus alors qu'une question de patience. (D^n est la matrice diagonale dont les éléments de la diagonale sont, dans l'ordre, $1, (1-4p)^n, (2p-1)^n, (2p-1)^n$.)

On est amené à calculer la puissance n-ème de certaines matrices, par exemple en utilisant la méthode donnée ci-dessus, pour une utilisation en probabilités dont nous donnons les grandes lignes. Le vocabulaire est donné à titre d'information, aucune connaissance spécifique n'est exigible.

• Un mobile se déplace aléatoirement (*marche aléatoire*). À chaque instant n ($n \in \mathbf{N}$), sa position X_n est une variable aléatoire qui prend ses valeurs dans l'ensemble $\{1, \dots, r\}$ (r entier naturel fixé $\geqslant 2$).

• Pour tout i, j appartenant à $\{1, \dots, r\}$, on suppose que la *probabilité de transition* $p_{i,j} = \mathrm{P}_{(X_n=i)}(X_{n+1} = j)$ ne dépend pas de n. On appelle *vecteur d'état à l'instant n* la matrice colonne C_n dont les termes successifs sont les probabilités $\mathrm{P}(X_n = 1), \dots, \mathrm{P}(X_n = r)$.

• Soit la matrice $A = \left(p_{i,j}\right)_{\substack{1\leqslant i\leqslant r \\ 1\leqslant j\leqslant r}}$, appelée *matrice de transition*.

La formule des probabilités totales, appliquée aux événements $\mathrm{P}(X_{n+1} = 1), \dots, \mathrm{P}(X_{n+1} = r)$, avec le sce $(X_n = i)_{1\leqslant i\leqslant r}$, conduit à :

$$X_{n+1} = AX_n$$

• On en déduit alors, par récurrence : $\forall n \in \mathbf{N}, C_n = A^n C_0$. Le calcul de A^n et la donnée de C_0 conduisent alors à la connaissance de C_n. Un éventuel passage à la limite permet de prévoir ce qui se passe pour les « grandes » valeurs de n.

Notons qu'il n'est pas toujours nécessaire d'expliciter A^n (voir l'exemple), et que la matrice de transition définie ici est la transposée de la matrice de transition telle qu'elle est définie dans l'enseignement de spécialité « pratique des graphes » de terminale ES.

On vérifie que la matrice A de l'exemple ci-dessus est la matrice de transition pour la marche aléatoire ainsi définie :

« Les sommets d'un carré sont numérotés 1, 2, 3 et 4 de telle façon que les côtés du carré relient le sommet 1 au sommet 2, le sommet 2 au sommet 3, le sommet 3 au sommet 4, le sommet 4 au sommet 1, les diagonales reliant elles le sommet 1 au sommet 3 ainsi que le sommet 2 au sommet 4.

Un pion se déplace sur les sommets de ce carré selon le protocole suivant :

• Le pion est sur le sommet 1 au départ.

• Lorsque le pion est à un instant donné sur un sommet du carré, il se déplace à l'instant suivant vers un sommet voisin (relié par un côté) avec la probabilité p ou vers un sommet opposé (relié par une diagonale) avec la probabilité $1 - 2p$. »

On a donc

$$C_n = A^n C_0 = \frac{1}{4} PD^n PC_0$$

Le pion est sur le sommet 1 au départ, donc $C_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, donc PC_0 est la première colonne de P, puis $D^n PC_0 = \begin{pmatrix} 1 \\ (1-4p)^n \\ (2p-1)^n \\ (2p-1)^n \end{pmatrix}$, puis

$$C_n = \frac{1}{4} P \begin{pmatrix} 1 \\ (1-4p)^n \\ (2p-1)^n \\ (2p-1)^n \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ (1-4p)^n \\ (2p-1)^n \\ (2p-1)^n \end{pmatrix}$$

En explicitant les probabilités $\mathrm{P}(X_n = i)$, on s'aperçoit que leur limite quand n tend vers $+\infty$ est égale à $1/4$. En effet, $p \in \left]0, 1/2\right[$, donc $-1 < 2p - 1 < 0$ et $-1 < 4p - 1 < 1$, donc $(1-4p)^n$ et $(2p-1)^n$ tendent vers 0 quand n tend vers $+\infty$. La suite de va (X_n) converge donc en loi vers une va X de loi uniforme sur $\{1, 2, 3, 4\}$.

Puissances d'une matrice, utilisation pour l'étude d'une suite

• Calcul de A^n. Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique $\mathcal{B}$ de $\mathbf{R}^3$ est $A = \begin{pmatrix} 5 & -8 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, et soit

$$v_1 = (1, 1, 1)\,,\ v_2 = (4, 2, 1)\,,\ v_3 = (4, 1, 0).$$

On vérifie que $f(v_1) = v_1, f(v_2) = 2v_2, f(v_3) = v_2 + 2v_3$, et que la famille $\mathcal{C} = (v_1, v_2, v_3)$ est une base de $\mathbf{R}^3$.

D'après la théorie du changement de base, on a alors $A = PTP^{-1}$, avec $T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, P = \begin{pmatrix} 1 & 4 & 4 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

On montre alors, par récurrence : $T^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix}$, ce qui permet le calcul de $A^n = \left(PTP^{-1}\right)^n = PT^nP^{-1}$.

• Utilisation pour l'étude d'une suite. Soit (u_n) la suite définie par :

$$\begin{cases} u_0 = 1\ ;\ u_1 = -1\ ;\ u_2 = 1 \\ \forall n \in \mathbf{N},\quad u_{n+3} = 5u_{n+2} - 8u_{n+1} + 4u_n \end{cases}$$

Avec $Y_n = \begin{pmatrix} u_{n+2} \\ u_{n+1} \\ u_n \end{pmatrix}$, on vérifie que, pour tout n, $Y_{n+1} = AY_n$. On en déduit, par récurrence, $Y_n = A^n Y_0$. On a donc $Y_n = PT^nP^{-1}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.
On achève les calculs, et la troisième ligne de la matrice obtenue donne, pour tout $n \in \mathbf{N}$: $u_n = 9 - 8 \cdot 2^n + 6n \cdot 2^{n-1}$.

Équation matricielle

Soit à résoudre un problème du type $\sum_{k=0}^{n} a_k X^k = A$, où la matrice $A = PDP^{-1}$ est diagonalisable, et X est une matrice à déterminer. L'idée générale est la suivante : on pose $Y = P^{-1}XP$. On a alors $X = PYP^{-1}$ et l'équation proposée est équivalente à

$$\sum_{k=0}^{n} a_k \left(PYP^{-1}\right)^k = PDP^{-1},$$

puis $P\left(\sum_{k=0}^{n} a_k Y^k\right) P^{-1} = PDP^{-1}$, puis $\sum_{k=0}^{n} a_k Y^k = D$. Le problème obtenu est plus simple à résoudre, car D est diagonale.

Soit $A = \begin{pmatrix} 16 & 4 & -4 \\ -18 & -4 & 5 \\ 30 & 8 & -7 \end{pmatrix}$. A est diagonalisable, on a $A = PDP^{-1}$, avec $P = \begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 1 \\ 2 & 1 & -2 \end{pmatrix}$ et $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

• On étudie l'équation $X^2 = A$, $X \in \mathcal{M}_3(\mathbf{R})$.. Soit $Y = P^{-1}XP$.
(**Analyse**) Si $X^2 = A$, alors, puisque $X = PYP^{-1}$, on a

$$\left(PYP^{-1}\right)^2 = PDP^{-1} \quad ; \quad PY^2P^{-1} = PDP^{-1} \quad ; \quad Y^2 = D$$

L'équation $Y^2 = D$ n'est pas simple à résoudre directement, mais on remarque que si $Y^2 = D$, alors $YD = DY$. En effet $YD = YY^2 = Y^3$, et $DY = Y^2Y = Y^3$.
On résout l'équation $YD = DY$, en cherchant Y sous la forme $Y = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}$. On trouve (facilement) que Y est une matrice diagonale. L'équation $Y^2 = D$ fournit alors $Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & 2\gamma' \end{pmatrix}$, avec

$\gamma, \gamma' \in \{-1, 1\}$. On a obtenu : si $X^2 = A$, alors $X = PYP^{-1}$, avec Y de la forme indiquée.

(**Synthèse**) Dans le raisonnement précédent, on a perdu l'équivalence. Il importe donc de vérifier que les matrices trouvées vérifient l'égalité $X^2 = A$, ce qui est bien le cas :

$$X^2 = \left(PYP^{-1}\right)^2 = PY^2P^{-1} = PDP^{-1} = A$$

• On étudie l'équation $MA = AM$, $M \in \mathcal{M}_3(\mathbf{R})$. Soit $Y = P^{-1}MP$.

On a $M = PYP^{-1}$. L'équation proposée est équivalente à :

$$PDP^{-1}PYP^{-1} = PYP^{-1}PDP^{-1} \;; \quad PDYP^{-1} = PYDP^{-1} \;;$$

$$DY = YD \;; \quad Y = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} = aE_{1,1} + bE_{2,2} + cE_{3,3}$$

avec $E_{i,j}$ la matrice dont tous les termes sont nuls sauf le terme en i-ème ligne et j-ème colonne qui est égal à 1.

M commute donc avec A ssi

$$M = PYP^{-1} = P\left(aE_{1,1} + bE_{2,2} + cE_{3,3}\right)P^{-1} = aM_1 + bM_2 + cM_3,$$

avec $M_i = PE_{i,i}P^{-1}$. L'ensemble des M qui commutent avec A est donc un espace vectoriel, le sous-espace vectoriel F de $\mathcal{M}_3(\mathbf{R})$ engendré par la famille (M_1, M_2, M_3), qui est une famille libre ($aM_1 + bM_2 + cM_3 = O$ implique $M = O$, puis $Y = O$, puis $a = b = c = 0$). Cette famille est donc une base de F, qui est donc de dimension 3.



Partie 3

Probabilités



Probabilité sur un ensemble fini

7

1. Espaces probabilisés finis

1.1 Définitions

Espace probabilisable fini

• Une **expérience aléatoire** est une expérience dont le résultat dépend du hasard.

• L'ensemble (non vide) des résultats élémentaires de l'expérience, appelé « univers des possibles », se note en général Ω.

Si Ω est fini, tout sous-ensemble, ou partie, A, de Ω est appelé **événement.**

• Avec Ω fini, le couple $(\Omega, \mathcal{P}(\Omega))$ est un **espace probabilisable fini**.

• L'événement « A **ou** B » est la réunion $A \cup B$ des événements A, B :

$$A \cup B = \{\omega \ ; \omega \in A \text{ ou } \omega \in B\}$$

• L'événement « A **et** B » est l'intersection $A \cap B$ des événements A, B :

$$A \cap B = \{\omega \ ; \omega \in A \text{ et } \omega \in B\}$$

• L'**événement contraire** de l'événement A est $\overline{A}$, il est défini par :

$$\overline{A} = \{\omega \ ; \omega \in \Omega \text{ et } \omega \notin A\}$$

• Les événements A, B sont dits **incompatibles** ssi leur intersection est l'ensemble vide :

$$A \cap B = \varnothing$$

• $\varnothing$ est l'**événement impossible**.

• Ω est l'**événement certain**.

On lance un dé à jouer, dont les faces sont numérotées de 1 à 6.
L'ensemble des résultats élémentaires de l'expérience est

$$\Omega = \{1\,;2\,;3\,;4\,;5\,;6\}$$

L'ensemble $\mathcal{P}(\Omega)$ est de cardinal $2^6 = 64$.
Événement A : « obtenir un nombre pair » :

$$A = \{2\,;4\,;6\}$$

Événement B : « obtenir un nombre premier » :

$$B = \{2\,;3\,;5\}$$

Événement A ou B : « obtenir un nombre pair ou premier » :

$$A \cup B = \{2\,;3\,;4\,;5\,;6\}$$

Événement A et B : « obtenir un nombre pair et premier » :

$$A \cap B = \{2\}$$

Événement contraire de A : « obtenir un nombre impair » :

$$\overline{A} = \{1\,;3\,;5\}$$

Les événements A et $\{1\}$: « obtenir la face n°1 » sont incompatibles.

Probabilité

Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable fini. Une **probabilité** sur Ω est une application $\mathrm{P} : \mathcal{P}(\Omega) \to [0, 1]\ ;\ A \mapsto \mathrm{P}(A)$ telle que :

- $\mathrm{P}(\Omega) = 1$;
- si A, B sont deux événements incompatibles ($A \cap B = \varnothing$), alors (propriété d'**additivité**) :

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$$

Le triplet $(\Omega, \mathcal{P}(\Omega), \mathrm{P})$ est alors un **espace probabilisé fini**.

1.2 Propriétés

- $\mathrm{P}(\varnothing) = 0$;
- $\mathrm{P}\left(\overline{A}\right) = 1 - \mathrm{P}(A)$;
- **Croissance**

Si $A \subset B$, alors

$$\mathrm{P}(A) \leqslant \mathrm{P}(B) \text{ et } \mathrm{P}\left(B \setminus A\right) = \mathrm{P}(B) - \mathrm{P}(A)$$

• **Additivité**
Si $A_1, \ldots, A_n$ sont des événements deux à deux incompatibles, alors

$$\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathrm{P}\left(A_k\right)$$

• **Formule du crible**, ou **Formule de Poincaré** pour deux événements A, B :

$$\mathrm{P}\left(A \cup B\right) = \mathrm{P}\left(A\right) + \mathrm{P}\left(B\right) - \mathrm{P}\left(A \cap B\right)$$

La formuledu crible se généralise :
– pour trois événements A_1, A_2, A_3 :

$$\begin{aligned}\mathrm{P}\left(A_1 \cap A_2 \cap A_3\right) = \mathrm{P}\left(A_1\right) + \mathrm{P}\left(A_2\right) + \mathrm{P}\left(A_3\right) \\ - \mathrm{P}\left(A_1 \cap A_2\right) - \mathrm{P}\left(A_1 \cap A_3\right) - \mathrm{P}\left(A_2 \cap A_3\right) \\ + \mathrm{P}\left(A_1 \cap A_2 \cap A_3\right)\end{aligned}$$

– pour n événements $A_1, \ldots, A_n$, $n \geqslant 2$:

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \left(\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right)\right)$$

La propriété d'additivité est d'usage constant. Pour déterminer la probabilité d'un événement, on le décompose en réunion d'événements plus simples, et deux à deux incompatibles.

Vous écrirez par extrapolation la formule du crible pour 4 événements. La formule générale n'a pas à être mémorisée, l'énoncé devrait vous la rappeler le cas échéant.

1.3 Exemples

Probabilité uniforme

Avec Ω fini non vide, l'application P définie sur $\mathcal{P}(\Omega)$ par

$$\mathrm{P}\left(A\right) = \frac{\mathrm{Card}\left(A\right)}{\mathrm{Card}\left(\Omega\right)} \left(= \text{« } \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} \text{ »}\right)$$

est une probabilité, la **probabilité uniforme** sur Ω.

On dit aussi qu'on est en situation d'**équiprobabilité**.

En situation d'équiprobabilité, on utilisera les formules de dénombrement vues au § 4 de l'Introduction, en commençant par déterminer $\mathrm{Card}\left(\Omega\right)$.

Dans le cas de tirages dans une urne, modèle auquel on peut toujours se ramener dans le cas où Ω est fini, on a :

- Tirage une à une et avec remise de k boules d'une urne de taille n :

$$\text{Card}(\Omega) = n^k$$

- Tirage une à une et sans remise de k boules d'une urne de taille n :

$$\text{Card}(\Omega) = n(n-1)\dots(n-k+1)$$

- Tirage exhaustif des n boules de l'urne :

$$\text{Card}(\Omega) = n!$$

- Tirage simultané de k boules de l'urne de taille n :

$$\text{Card}(\Omega) = \binom{n}{k}$$

Tirage simultané et tirage une à une sont identiques si on s'intéresse uniquement au résultat obtenu, et pas à l'ordre d'apparition des boules.

Dans le cas très fréquent de tirage avec remise, vous devez penser à l'indépendance des tirages (§ 7.1.6) ! Et à la loi binomiale, § 7.4.2.

Probabilité définie à partir des événements élémentaires

Si $\Omega = \{\omega_1, \dots, \omega_n\}$, la donnée de n nombres réels $p_1, \dots, p_n$ tels que

- $\forall i \in [\![1, n]\!],\ p_i \geqslant 0$;
- $\displaystyle\sum_{i=1}^{n} p_i = 1$

permet de définir une probabilité en posant, pour tout $A \in \mathcal{P}(\Omega)$:

$$\text{P}(A) = \sum_{\{k\,;\, w_k \in A\}} p_k$$

On a $p_k = \text{P}\left(\{\omega_k\}\right)$. Avec $p_k = \frac{1}{\text{Card}(\Omega)}$, on retrouve la probabilité uniforme.

1.4 Probabilité conditionnelle

Définition. Soit $(\Omega, \mathcal{P}(\Omega), \text{P})$ un espace probabilisé fini, et $B \subset \Omega$ tel que $\text{P}(B) \neq 0$. La **probabilité conditionnelle sachant** B est l'application

$$\text{P} : \mathcal{P}(\Omega) \to [0\ ;\ 1] \quad ; \quad A \mapsto \text{P}_B(A) = \frac{\text{P}(A \cap B)}{\text{P}(B)}$$

On vérifie que P_B est une probabilité.

On rencontre la notation $\mathrm{P}_B(A) = \mathrm{P}\left(A/B\right)$, dangereuse car elle laisse penser que A/B serait un événement.

Utilisation

Formule des probabilités composées

- Si A est de probabilité non nulle :

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}_A(B)$$

- Si $A_1 \cap A_2 \cap \cdots \cap A_{n-1}$ est de probabilité non nulle :

$$\mathrm{P}(A_1 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\,\mathrm{P}_{A_1}(A_2)\cdots\mathrm{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n)$$

• Un sac contient 5 jetons blancs et 4 jetons noirs. On tire au hasard et sans remise 3 jetons du sac. La probabilité d'obtenir un jeton blanc, puis un noir, puis un blanc, est, avec des notations évidentes :

$$\begin{aligned}\mathrm{P}(B_1 \cap N_2 \cap B_3) &= \mathrm{P}(B_1)\,\mathrm{P}_{B_1}(N_2)\,\mathrm{P}_{B_1\cap N_2}(B_3)\\ &= \frac{5}{9}\times\frac{4}{8}\times\frac{4}{7} = \frac{10}{63}\end{aligned}$$

résultat qu'on peut établir directement par équiliprobabilité, en calculant $\mathrm{Card}(\Omega) = 9\cdot 8\cdot 7$, puis $\mathrm{Card}(B_1 \cap N_2 \cap B_3) = 5\cdot 4\cdot 4$.

• Dans un jeu vidéo, la probabilité de franchir le niveau k (événement N_k) quand on a franchi le niveau $k-1$ est $\frac{1}{k}$, pour $k \geqslant 1$. On part du niveau 0, et le jeu comporte n niveaux. La probabilité d'arriver au bout du jeu est :

$$\begin{aligned}\mathrm{P}(N_1 \cap \cdots \cap N_n) &= \mathrm{P}(N_1)\,\mathrm{P}_{N_1}(N_2)\cdots\mathrm{P}_{N_1\cap\cdots\cap N_{n-1}}(N_n)\\ &= 1\cdot\frac{1}{2}\cdot\ldots\cdot\frac{1}{n} = \frac{1}{n!}\end{aligned}$$

1.5 Formule des probabilités totales

Définition. Soit $(\Omega, \mathcal{P}(\Omega), \mathrm{P})$ un espace probabilisé fini, I un ensemble fini d'indices.

On dit que$(A_i)_{i\in I}$ est un **système complet d'événements** (sce) ssi :

- $\forall i \in I, \quad \mathrm{P}(A_i) \neq 0$;
- $\forall i, j \in I$, si $i \neq j, \quad A_i \cap A_j = \varnothing$;
- $\displaystyle\bigcup_{i\in I} A_i = \Omega$.

Avec X v.a tel que $X(\Omega) = \{x_1, \ldots, x_n\}$, les événements

$$(X = x_1), (X = x_2), \ldots, (X = x_n)\}$$

constituent un sce. Voir § 7.2.

Formule des probabilités totales

Soit $(A_i)_{i \in I}$, un sce de Ω. Alors, pour tout événement B de Ω :

$$\mathrm{P}(B) = \sum_{i \in I} \mathrm{P}(B \cap A_i) = \sum_{i \in I} \mathrm{P}_{A_i}(B)\,\mathrm{P}(A_i)$$

Formule de Bayes

Avec A_{i_0} un des éléments du sce $(A_i)_{i \in I}$:

$$\mathrm{P}_B\left(A_{i_0}\right) = \frac{\mathrm{P}\left(A_{i_0} \cap B\right)}{\mathrm{P}(B)} = \frac{\mathrm{P}_{A_{i_0}}(B)\,\mathrm{P}\left(A_{i_0}\right)}{\sum_{i \in I} \mathrm{P}_{A_i}(B)\,\mathrm{P}(A_i)}$$

Démonstration. Formule des probabilités totales : B est la réunion des événements deux à deux incompatibles $B \cap A_i$. L'additivité de la probabilité donne le premier résultat. En appliquant la formule des probabilités composées à chacune des probabilités $\mathrm{P}(B \cap A_i)$, on obtient le deuxième résultat.

Formule de Bayes : définition de la probabilité conditionnelle, puis formule des probabilités composées pour le numérateur, et formule des probabilités totales pour le dénominateur. Remarquez que le numérateur est un des termes de la somme du dénominateur.

La formule de Bayes doit être considérée comme une conséquence facile de la formule de probabilités totales. Pour appliquer celle-ci, l'énoncé doit fournir (par les conditions de l'expérience aléatoire) les probabilités $\mathrm{P}_{A_i}(B)$, $\mathrm{P}(A_i)$, et vous devez reconnaître le système complet d'événements en jeu.

Un sac contient n jetons numérotés de 1 à n. On tire au hasard un jeton. Si on a obtenu le jeton n°i, on lance i fois une pièce de monnaie équilibrée. On cherche la probabilité de l'événement B : « on obtient 0 pile ».

Le sce est $(A_i)_{1 \leqslant i \leqslant n}$, avec A_i: « obtenir le jeton n°i » ; en effet, quelle que soit l'issue de l'expérience, un et un seul des événements A_i se

produit. L'énoncé fournit

$$\forall i \in [\![1, n]\!], \qquad \mathrm{P}(A_i) = \frac{1}{n} \quad ; \quad \mathrm{P}_{A_i}(B) = \left(\frac{1}{2}\right)^i$$

La formule des probabilités totales donne alors

$$\mathrm{P}(B) = \sum_{i=1}^{n} \frac{1}{n} \times \left(\frac{1}{2}\right)^i = \frac{1}{n} \times \left(1 - \left(\frac{1}{2}\right)^n\right)$$

en utilisant l'identité géométrique.
La formule de Bayes fournit alors, par exemple :

$$\mathrm{P}_{\mathrm{B}}(A_1) = \frac{\mathrm{P}(A_1 \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}_{A_1}(B)\,\mathrm{P}(A_1)}{\mathrm{P}(B)} = \frac{1}{2\left(1 - \left(\frac{1}{2}\right)^n\right)}$$

1.6 Indépendance

Définition

• Deux événements A, B sont dits **indépendants** ssi

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$$

ce qui est équivalent à $\mathrm{P}_B(A) = \mathrm{P}(A)$ si B est de probabilité non nulle.

• Les événements $A_1, \dots, A_n$ sont dits **indépendants**, ou **mutuellement indépendants**, ssi

$$\forall I \subset [\![1, n]\!], \quad \mathrm{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathrm{P}(A_i)$$

Exemples fondamentaux

• Les résultats successifs du jet d'une pièce de monnaie, d'un dé, etc. sont des événements indépendants.

• Les résultats successifs de tirages AVEC REMISE d'une boule dans une urne sont des événements indépendants.

Propriétés

• Si les événements $A_1, \ldots, A_n$ sont indépendants, alors il en est de même des événements $B_1, \ldots, B_n$, avec $B_i = A_i$ ou $B_i = \overline{A_i}$.

• Les événements $\varnothing$ et Ω sont indépendants de tout événement A.

• Si les événements $A_1, \ldots, A_n$ sont indépendants, alors ils sont deux à deux indépendants. La réciproque est fausse.

Ne confondez pas indépendance et incompatibilité. L'indépendance est une notion probabiliste, pas l'incompatibilité. Si deux événements sont incompatibles, ils ne sont pas indépendants, et s'ils sont indépendants, ils ne sont pas incompatibles, à moins que l'un des deux ne soit de probabilité nulle.

2. Variables aléatoires sur un ensemble fini

2.1 Définitions

• Soit $\Omega, \mathcal{P}(\Omega), \mathrm{P})$ un espace probabilisé fini. Une **variable aléatoire** est une application de Ω dans $\mathbf{R}$.

Une variable aléatoire (v.a) est donc une observation numérique sur le résultat d'une expérience aléatoire.

On parlera au besoin d'une v.a **finie** pour préciser que Ω est fini.

• Pour tout $x \in \mathbf{R}$, l'événement

$$X^{-1}(x) = \{\omega \ ; \omega \in \Omega, X(\omega) = x\}$$

est noté, de façon plus parlante, $(X = x)$.

• On note de même $(X \leqslant x)$, $(a \leqslant X \leqslant b)\ldots$

• $X(\Omega)$ est l'ensemble des valeurs prises par X.

• Déterminer la **loi de probabilité** de X, c'est déterminer $X(\Omega)$, et, pour chaque $x_i \in X(\Omega)$, déterminer la probabilité $\mathrm{P}(X = x_i)$.

• Réciproquement, soit $\{(x_i, p_i) \ ; i \in I\}$ un ensemble fini de couples de nombres réels. Pour vérifier que cet ensemble est la loi de probabilité d'une v.a finie X, avec $\mathrm{P}(X = x_i) = p_i$, il suffit de vérifier :

– $\forall i \in I, \ p_i \geqslant 0$;

– $\displaystyle\sum_{i \in I} p_i = 1$.

2.2 Propriétés

• La somme, le produit, le quotient (si le dénominateur ne s'annule pas) de deux v.a définies sur le même ensemble Ω est une v.a.

• Si ϕ est une fonction définie sur $X(\Omega)$, alors la composée $\phi \circ X$ est une v.a définie sur Ω.

• L'ensemble des v.a définies sur Ω est un ev pour les opérations $X + Y$(somme de deux v.a), αX (produit d'une v.a par un nombre réel). Voir § 5.1.2.

2.3 Indépendance de variables aléatoires

Définitions

• Deux va finies X, Y sont dites **indépendantes** ssi :

$\forall x \in X(\Omega), \forall y \in Y(\Omega),$

$$\mathrm{P}((X = x) \cap (Y = y)) = \mathrm{P}(X = x)\,\mathrm{P}(Y = y)$$

• Les v.a $X_1, X_2, \cdots, X_n$ sont dites **indépendantes** (ou **mutuellement indépendantes**) ssi :

$\forall x_1 \in X_1(\Omega), \forall x_2 \in X_2(\Omega), \cdots, \forall x_n \in X_n(\Omega),$

$$\mathrm{P}\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = \prod_{i=1}^{n} \mathrm{P}(X_i = x_i)$$

Propriétés

• Si les v.a $X_1, \cdots, X_n$ sont indépendantes, alors elles sont indépendantes deux à deux. La réciproque est fausse.

• Si $X_1, \cdots, X_n$ sont indépendantes, alors toute fonction de $X_1, \cdots, X_p$ est indépendante de toute fonction de $X_{p+1}, \cdots, X_n$, $1 \leqslant p < n$. En particulier, si X, Y sont indépendantes, alors toute fonction de X est indépendante de toute fonction de Y.

2.4 Espérance, ou moyenne

Définition. Soit X une v.a définie sur Ω fini. L'**espérance**, ou **moyenne**, de X est le nombre réel noté $\mathrm{E}(X)$ et défini par

$$\mathrm{E}(X) = \sum_{x_i \in X(\Omega)} x_i \mathrm{P}(X = x_i)$$

Cette définition est à rapprocher de la définition de la moyenne d'une série statistique : la fréquence de la valeur observée x_i de la variable statistique X est remplacée par la probabilité de la valeur x_i pour la v.a X.

Un cas particulier d'utilisation fréquente :

$$\text{Si } X(\Omega) \subset [\![0, n]\!], \quad \mathrm{E}(X) = \sum_{k=0}^{n} k\mathrm{P}(X = k)$$

Propriétés

- **Linéarité de l'espérance**

Soit X et Y deux v.a définies sur Ω fini et a, b deux nombres réels. Alors

$$\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$$

$$\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$$

L'espérance est donc une forme linéaire sur l'espace vectoriel des v.a définies sur Ω fini, voir chapitre 5.

- **Théorème de transfert**

Soit X une v.a définie sur Ω fini, et φ une application définie sur $X(\Omega)$. Alors

$$\mathrm{E}(\varphi(X)) = \sum_{x_i \in X(\Omega)} \varphi(x_i)\,\mathrm{P}(X = x_i)$$

En particulier, pour $r \in \mathbf{N}^*$,

$$\mathrm{E}(X^r) = \sum_{x_i \in X(\Omega)} x_i^r \mathrm{P}(X = x_i)$$

$\mathrm{m}_r(X) = \mathrm{E}(X^r)$ est appelé **moment d'ordre** r de X.

2.5 Variance, écart-type

Définitions

- soit X une v.a définie sur Ω fini. La **variance** de X est le nombre réel noté $\mathrm{V}(X)$ et défini par :

$$\mathrm{V}(X) = \sum_{x_i \in X(\Omega)} [x_i - \mathrm{E}(X)]^2 \mathrm{P}(X = x_i)$$

D'après le théorème de transfert, on a

$$\mathrm{V}(X) = \mathrm{E}\left([X - \mathrm{E}(X)]^2\right)$$

La variance est la moyenne des carrés des écarts à la moyenne.

• L'**écart-type** $\sigma(X)$ est la racine carrée de la variance :

$$\sigma(X) = \sqrt{\mathrm{V}(X)}$$

Propriétés

• Pour toute v.a X sur Ω, $\mathrm{V}(X) \geqslant 0$.
Si $\mathrm{V}(X) = 0$, alors X est une v.a certaine :

$$X(\Omega) = \{c\} \quad ; \quad \mathrm{P}(X = c) = 1$$

• Soit X une v.a finie, et a, b deux nombres réels. Alors

$$\mathrm{V}(aX + b) = a^2\mathrm{V}(X)$$

Démonstration : Avec $Y = aX + b$, on a $Y - \mathrm{E}(Y) = a(X - \mathrm{E}(X))$, donc $(Y - \mathrm{E}(Y))^2 = a^2(X - \mathrm{E}(X))^2$, donc

$$\mathrm{V}(aX + b) = \mathrm{V}(Y) = \mathrm{E}\left((Y - \mathrm{E}(Y))^2\right) = a^2\mathrm{E}\left((X - \mathrm{E}(X))^2\right)$$

d'où le résultat.

• **Théorème de Koenig-Huygens**
Soit X une v.a définie sur Ω fini.
Alors $\quad \mathrm{V}(X) = \sum_{x_i \in X(\Omega)} {x_i}^2\mathrm{P}(X = x_i) - [\mathrm{E}(X)]^2$

Démonstration : on développe $(x_i - \mathrm{E}(X))^2$, puis on utilise les règles de calcul sur le signe $\sum$.
D'après le théorème de transfert, on a donc :

$$\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2$$

La variance est égale à la moyenne des carrés diminuée du carré de la moyenne.

• **Variance d'une somme de 2 v.a finies**
Soit X et Y deux v.a définies sur Ω fini.

– Si X et Y sont indépendantes, alors

$$\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y)$$

– Dans le cas général :

$$V(X+Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y)$$

avec

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y),$$

avec

$$E(XY) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j P\left(X = x_i, Y = y_j\right)$$

• **Généralisation à *n* v.a finies**

Soit $X_1, \ldots, X_n$ n v.a définies sur Ω fini.

– Si les X_i sont deux à deux indépendantes, alors

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$$

– Dans le cas général

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{Cov}\left(X_i, Y_j\right)$$

Pour un exemple d'utilisation de ces deux dernières formules, voyez § 7.4.2 et 7.4.3

Attention :

$V(X-Y) = V(X) + V(Y) - 2\mathrm{Cov}(X, Y)$, voir § 7.3.3.

3. Couple de variables aléatoires finies

3.1 Définitions

Soit X, Y deux variables aléatoires finies.

• Déterminer la **loi de couple** (X, Y), c'est déterminer, pour chaque $x_i \in X(\Omega)$ et chaque $y_j \in Y(\Omega)$ la probabilité

$$P\left((X = x_i) \cap \left(y = y_j\right)\right)$$

On rencontre aussi les notations :

$$P\left((X = x_i) \text{ et } \left(y = y_j\right)\right) \quad ; \quad P\left((X = x_i), \left(y = y_j\right)\right)$$

• La connaissance de la loi du couple(X, Y) permet de trouver les **lois marginales**, c'est-à-dire les lois de X et de Y, en utilisant la formule de

probabilités totales. Avec le sce $\left(Y = y_j\right)_{y_j \in Y(\Omega)}$, on a

$$\forall x_i \in X(\Omega), \quad P(X = x_i) = \sum_{x_i \in X(\Omega)} P\left((X = x_i) \cap \left(Y = y_j\right)\right)$$

et de même pour $P\left(Y = y_j\right)$, avec le sce $(X = x_i)_{x_i \in X(\Omega)}$.

• Pour y_j fixé, la **loi conditionnelle** de la variable X sachant $\left(Y = y_j\right)$, notée $X/\left(Y = y_j\right)$, est donnée par

$$P_{(Y=y_j)}(X = x_i) = \frac{P\left((X = x_i) \cap \left(Y = y_j\right)\right)}{P\left(Y = y_j\right)}$$

et de même pour la loi de la variable Y sachant $(X = x_i)$.

3.2 Théorème de transfert

Soit X, Y deux va finies, φ une fonction de deux variables définie sur $X(\Omega) \times Y(\Omega)$.

Alors

$$E(\varphi(X, Y)) = \sum_{(x_i, y_j) \in X(\Omega) \times Y(\Omega)} \varphi\left(x_i, y_j\right) P\left((X = x_i) \cap \left(Y = y_j\right)\right)$$

En particulier,

$$E(XY) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j P\left(X = x_i, Y = y_j\right)$$

3.3 Covariance

Soit X, Y deux v.a finies. Pour calculer $V(X + Y)$, on peut utiliser le théorème de Koenig-Huygens.

L'identité remarquable $(a + b)^2 = a^2 + 2ab + b^2$, et la linéarité de l'espérance fournissent alors :

$$V(X + Y) = V(X) + V(Y) + 2[E(XY) - E(X)E(Y)]$$

Ceci amène à poser la définition suivante.

Définition. Soit X, Y deux v.a finies. La covariance de (X, Y) est le nombre réel

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Propriétés

- $V(X+Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$.
- $\text{Cov}(X, Y)$ est de signe de quelconque.
- On a aussi $\text{Cov}(X, Y) = E([X - E(X)][Y - E(Y)])$.
- Si X, Y, Z sont des v.a sur Ω fini, et si a est un nombre réel, on a

– $\text{Cov}(X, X) = V(X)$;

– $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;

– $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$;

– $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$;

– $\text{Cov}(aX, Y) = \text{Cov}(X, aY) = a\,\text{Cov}(X, Y)$.

- La formule $V(X+Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$ permet le calcul de $\text{Cov}(X, Y)$, si les variances de X, Y et $X + Y$ sont connues :

$$\text{Cov}(X, Y) = \frac{1}{2}\left(V(X+Y) - V(X) - V(Y)\right)$$

- **Cas de deux variables indépendantes**

Soient X, Y deux va finies indépendantes. On a alors les propriétés suivantes, équivalents entre elles :

$$E(XY) = E(X)\,E(Y)$$

$$\text{Cov}(X, Y) = 0$$

$$V(X+Y) = V(X) + V(Y)$$

Les réciproques sont fausses. La covariance de deux v.a finies non indépendantes peut être égale à 0. On peut juste dire que si la covariance de (X, Y) est différente de 0, alors X, Y ne sont pas indépendantes.

3.4 Coefficient de corrélation linéaire

Définition. Soit X, Y deux va finies, non certaines. Le **coefficient de corrélation linéaire** de X, Y est le nombre réel $r_{X,Y}$ défini par

$$\boldsymbol{r_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma(X)\,\sigma(Y)}}$$

Propriétés

- $|r_{X,Y}| \leqslant 1$;
- si $|r_{X,Y}| = 1$, alors il existe a, b réels tels que $Y = aX + b$, et a est du signe de r.
- Si X, Y sont indépendantes, alors $r_{X,Y} = 0$. La réciproque est fausse.

Démonstration. Le polynôme en λ

$$P(\lambda) = \mathrm{V}(\lambda X + Y) = \lambda^2 \mathrm{V}(X) + 2\lambda \mathrm{Cov}(X, Y) + \mathrm{V}(Y)$$

est positif ou nul ; son discriminant est donc négatif ou nul ; on en déduit

$$[\mathrm{Cov}(X, Y)]^2 \leqslant \mathrm{V}(X)\,\mathrm{V}(Y)$$

d'où le premier résultat. Il y a égalité ssi le polynôme P admet une racine λ, ssi il existe c tel que $\lambda X + Y = c$, avec $\lambda = -\frac{\mathrm{Cov}(X,Y)}{\mathrm{V}(X)}$, d'où le deuxième résultat. Si X, Y sont indépendantes, alors $\mathrm{Cov}(X, Y) = 0$, d'où le troisième résultat.

4. Lois finies usuelles

4.1 Loi de Bernoulli

Soit $p \in]0\,;1[$. On dit que la v.a X suit la **loi de Bernoulli de paramètre** p, et on note $X \hookrightarrow \mathcal{B}(1, p)$, ou $X \hookrightarrow \mathcal{B}(p)$, ssi

$$X(\Omega) = \{0\,;1\} \quad ; \quad \mathrm{P}(X = 1) = p \quad ; \quad \mathrm{P}(X = 0) = 1 - p$$

Espérance, variance

$$\mathrm{E}(X) = p \quad ; \quad \mathrm{V}(X) = p(1 - p)$$

On dit que la v.a X est une **variable de Bernoulli** ssi $X(\Omega) = \{0\,;1\}$. X suit alors la loi $\mathcal{B}(1, p)$, avec $p = \mathrm{P}(X = 1)$.

La **variable indicatrice** 1_A de l'événement A (ni certain, ni impossible), définie par

$$1_A = \begin{cases} 1 & \text{si } A \text{ est réalisé} \\ 0 & \text{sinon} \end{cases}$$

est une variable de Bernoulli de paramètre $\mathrm{P}(A)$.

Une **épreuve de Bernoulli** est une expérience aléatoire ayant deux issues possibles : le succès et l'échec. La variable indicatrice du succès est alors une variable de Bernoulli de paramètre la probabilité du succès.

Simulation informatique. On utilise la fonction random, qui retourne un nombre aléatoire suivant la loi uniforme sur $[0\,;1]$: si X suit une telle loi et $x \in [0\,;1]$, alors $\mathrm{P}(X \leqslant x) = x$ (voir § 9.2.1).

```
function bernoull(  :real) :integer ;
begin
if random<=  then bernoull:=1 else bernoull:=0;
end;
```

4.2 Loi binomiale

Soit $n \in \mathbf{N}^*, p \in]0\,;1[$. On dit que la v.a X suit la **loi binomiale** de paramètres n, p, et on note $X \hookrightarrow \mathcal{B}(n,p)$, ssi

$$X(\Omega) = [\![0,n]\!] \quad ; \quad \forall k \in [\![0,n]\!], \mathrm{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Modèle. X est le nombre de succès lors de n épreuves identiques et indépendantes, la probabilité du succès à chaque épreuve étant p.

Espérance, variance

$$\mathrm{E}(X) = np \quad ; \quad \mathrm{V}(X) = np(1-p)$$

Loi binomiale et loi de Bernoulli. $X \hookrightarrow \mathcal{B}(n,p)$ ssi X est la somme de n variables de Bernoulli indépendantes et de même loi $\mathcal{B}(1,p)$.

Stabilité. Si $X \hookrightarrow \mathcal{B}(n,p)$, $Y \hookrightarrow \mathcal{B}(m,p)$, X et Y indépendantes, alors

$$X + Y \hookrightarrow \mathcal{B}(n+m,p)$$

• L'obtention de la loi de X à partir du modèle est relativement difficile : $(X=k)$ est la réunion de $\binom{n}{k}$ événements incompatibles : chacun de ces événements est une suite de n épreuves parmi lesquelles il y a k succès ; il y a $\binom{n}{k}$ telles suites, autant que de manières de choisir les k épreuves parmi n qui amènent le succès ; et chacune de ces suites est de probabilité $p^k(1-p)^{n-k}$, par indépendance des épreuves.

D'autre part les situations qui amènent à une loi binomiale sont très fréquentes. C'est pourquoi il convient d'avoir à l'esprit et la loi et le modèle.

• Démontrons le résultat sur l'espérance (on note $q = 1-p$) :

$$\begin{aligned}\mathrm{E}(X) &= \sum_{k=0}^{n} k\mathrm{P}(X=k) = \sum_{k=1}^{n} k\binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=1}^{n} n\binom{n-1}{k-1} p^k q^{n-k} = n\sum_{i=0}^{n-1}\binom{n-1}{i} p^{i+1} q^{n-1-i} \\ &= np\sum_{i=0}^{n-1}\binom{n-1}{i} p^i q^{n-1-i} = np(p+q)^{n-1} = np\end{aligned}$$

Étapes du calcul : définition de $\mathrm{E}(X)$; la somme démarre à $k = 1$; recours aux factorielles ; mise en facteur de n et changement de variable $i = k - 1$; mise en facteur de p ; formule du binôme.

• Pour le calcul de $\mathrm{V}(X)$, on écrit :

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2 = \mathrm{E}(X(X-1)+X) - [\mathrm{E}(X)]^2 \\ &= \mathrm{E}(X(X-1)) + \mathrm{E}(X) - [\mathrm{E}(X)]^2\end{aligned}$$

et on calcule $\mathrm{E}(X(X-1))$ par le théorème de transfert, de façon analogue au calcul de $\mathrm{E}(X)$.

• $X \hookrightarrow \mathcal{B}(n,p)$ ssi X est la somme de n variables de Bernoulli indépendantes et de même loi $\mathcal{B}(1,p)$. En effet, si $X \hookrightarrow \mathcal{B}(n,p)$, alors $X = \sum_{i=1}^{n} X_i$ avec $X_i = \begin{cases} 1 & \text{si succès à la } i^{\text{ème}} \text{ épreuve} \\ 0 & \text{sinon} \end{cases}$

et les X_i sont indépendantes car les épreuves le sont.

Réciproquement, si X est la somme de n variables de Bernoulli indépendantes de même paramètre p, alors X compte le nombre de succès $(X_i = 1)$ lors de n épreuves identiques et indépendantes, la probabilité du succès à chaque épreuve étant p, et donc $X \hookrightarrow \mathcal{B}(n,p)$.

• On utilise cette propriété pour calculer facilement $\mathrm{E}(X)$ et $\mathrm{V}(X)$ avec $X \hookrightarrow \mathcal{B}(n,p)$:

– d'après la linéarité de l'espérance :

$$E(X) = \mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}(X_i) = \sum_{i=1}^{n} p = np$$

– les v.a X_i sont indépendantes, donc :

$$\mathrm{V}(X) = \mathrm{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{V}(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

• Démontrons le résultat sur la stabilité de la loi binomiale, ce qui est une occasion de déterminer la loi d'une somme.

Soit $X \hookrightarrow \mathcal{B}(n,p)$, $Y \hookrightarrow \mathcal{B}(m,p)$, X et Y indépendantes, et $Z = X + Y$.

Z prend ses valeurs dans $[\![0, n+m]\!]$, et, pour tout $k \in [\![0, n+m]\!]$,

$$(Z = k) = \bigcup_{i=0}^{k} [(X = i) \cap (Y = k - i)]$$

Par additivité de la probabilité, puis indépendance de X, Y, on a

$$\begin{aligned}\mathrm{P}(Z=k) &= \sum_{i=0}^{n} \mathrm{P}(X=i)\,\mathrm{P}(Y=k-i) \\ &= \sum_{i=0}^{n} \binom{n}{i}\binom{m}{k-i} p^k q^{n+m-k}\end{aligned}$$

d'où le résultat, par la formule de Vandermonde (§ 4.3 de l'Introduction et § 7.4.3).

• La stabilité de la loi binomiale se généralise : si les $(X_i)_{1\leqslant i\leqslant m}$ sont indépendantes et telles que $X_i \hookrightarrow \mathcal{B}(n_i, p)$, alors

$$Y_m = \sum_{i=1}^{m} X_i \hookrightarrow \mathcal{B}\left(\sum_{i=1}^{m} n_i, p\right)$$

Démonstration par récurrence.

Simulation de $X \hookrightarrow \mathcal{B}(n, p)$.

```
function binom(n:integer; :real):real ;
var X,k:integer;
begin
 X:=0;  {X va compter le nombre de succès à l'épreuve random<  }
 for k:=1 to n do if random<  then X:=X+1;
 binom:=X;
end;
```

4.3 Loi hypergéométrique

Soit $n, N \in \mathbf{N}^*$ et $p \in]0\,;1[$ tels que $Np \in \mathbf{N}^*$. On dit que la v.a X suit la **loi hypergéométrique** de paramètres N, n, p, et on note $X \hookrightarrow \mathcal{H}(N, n, p)$, ssi

$$X(\Omega) \subset [\![0, n]\!]\,; \quad \forall k \in [\![0, n]\!], \quad \mathrm{P}(X=k) = \frac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}$$

avec la convention $\binom{n}{k} = 0$ si $k < 0$ ou $k > n$.

Modèle. On prélève au hasard un échantillon de n individus dans une population de taille N. Parmi ces N individus, une proportion p présente le caractère étudié C. X est le nombre d'individus de l'échantillon présentant le caractère étudié.

Espérance

$$\mathrm{E}(X) = np$$

• Appelons « individu-C » un individu présentant les caractère étudié. On trouve facilement la loi de X à partir de son modèle, par des considérations de dénombrement : on est en situation d'équiprobabilité, et il s'agit d'un tirage simultané de n individus parmi N, donc $\text{Card}\,(\Omega) = \binom{N}{n}$; on constitue un échantillon réalisant l'événement $(X = k)$ en choisissant. k individus-C parmi Np individus-C, puis $n - k$ individus-non-C parmi $N\,(1 - p)$ individus-non-C, ce qui donne $\text{Card}\,((X = k))$.

• $\mathcal{H}\,(N, n, p)$ est une loi de probabilité, par conséquent

$$\sum_{k=0}^{n} \binom{Np}{k} \binom{N\,(1-p)}{n-k} = \binom{N}{n}$$

ce qui démontre la formule de Vandermonde, voir l'Introduction § 4.3.

• X est la somme de Np variables de Bernoulli X_i, non indépendantes, avec $X_i = 1$ ssi le i^{me} individu-C est présent dans l'échantillon, 0 sinon.

X_i est une variable de Bernoulli de paramètre

$$\text{P}\,(X_i = 1) = \frac{1 \times \binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$$

En effet, une fois choisi le i^{me} individu-C, on complète l'échantillon en choisissant $n - 1$ individus parmi $N - 1$. On a donc

$$X = \sum_{i=1}^{Np} X_i \quad \text{avec} \quad X_i \hookrightarrow \mathcal{B}\left(\frac{n}{N}\right) ;$$

$$\text{E}\,(X_i) = \frac{n}{N} ; \quad \text{V}\,(X_i) = \frac{n}{N}\left(\frac{n}{N} - 1\right)$$

• Ceci permet le calcul de $\text{E}\,(X)$, en utilisant la linéarité de l'espérance. On trouve

$$\text{E}\,(X) = Np\frac{n}{N} = np$$

• Pour le calcul de $\text{V}\,(X)$, on utilise la formule du § 7.2.5 :

$$\text{V}\,(X) = \text{V}\left(\sum_{i=1}^{Np} X_i\right) = \sum_{i=1}^{Np} \text{V}\,(X_i) + 2 \sum_{1 \leqslant i < j \leqslant Np} \text{Cov}\left(X_i, Y_j\right)$$

avec $\text{Cov}\left(X_i, X_j\right) = \text{E}\left(X_i X_j\right) - \text{E}\,(X_i)\,\text{E}\left(X_j\right)$. Tout est connu, hormis $\text{E}\left(X_i X_j\right)$ et le nombre de termes dans la somme double.

– X_iX_j est une variable de Bernoulli de paramètre

$$E\left(X_iX_j\right) = P\left(X_iX_j = 1\right) = \frac{1 \times 1 \times \binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)}$$

– Tous les termes de la somme double sont égaux à

$$\operatorname{Cov}\left(X_i, X_j\right) = \frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2 = \frac{n(n-N)}{N(N-1)}$$

et il y a $\binom{Np}{2}$ tels termes. En effet, il a autant de couples (i, j) tels que $1 \leqslant i < j \leqslant Np$ qu'il y a de parties $\{i, j\}$ à deux éléments dans un ensemble à Np éléments.

– Il ne reste plus qu'à achever les calculs, on trouve

$$\mathrm{V}(X) = np(1-p)\frac{N-n}{N-1}$$

• Pour la simulation informatique de $X \hookrightarrow \mathcal{H}(N, n, p)$, voir § 10.2.3.

4.4 Loi uniforme sur $[\![1, n]\!]$

On dit que la v.a X suit la loi uniforme sur $[\![1, n]\!]$, et on note $X \hookrightarrow \mathcal{U}\left([\![1, n]\!]\right)$, ssi

$$X(\Omega) = [\![1, n]\!] \;\; ; \;\; \forall k \in [\![1, n]\!], \quad \mathrm{P}(X = k) = \frac{1}{n}$$

Espérance, variance

$$\mathrm{E}(X) = \frac{n+1}{2} \;\; ; \;\; \mathrm{V}(X) = \frac{n^2-1}{12}$$

• Pour le calcul de $\mathrm{E}(X)$, on utilise la formule $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$. Pour le calcul de $\mathrm{V}(X)$, on utilise le théorème de Koenig-Huygens, puis, pour le calcul de $\mathrm{E}\left(X^2\right)$, la formule $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

• Avec $a, b \in \mathbf{N}, a < b$, on dit que Y suit la loi uniforme sur $[\![a, b]\!]$ ssi, pour tout $k \in [\![a, b]\!]$, $\mathrm{P}(Y = k) = \frac{1}{n}$, avec $n = b - a + 1$, nombre d'éléments de l'ensemble $[\![a, b]\!]$. La v.a $X = Y - a + 1$ suit alors la loi

uniforme sur $[\![1, n]\!]$, car $1 \leqslant k \leqslant n \Leftrightarrow a \leqslant k + a - 1 \leqslant b$, et

$$\forall k \in [\![1, n]\!], \quad \mathrm{P}(X = k) = \mathrm{P}(Y = k + a - 1) = \frac{1}{n}$$

On a alors $Y = X + a - 1$, et des formules $\mathrm{E}(\alpha X + \beta) = \alpha \mathrm{E}(X) + \beta$ et $\mathrm{V}(\alpha X + \beta) = \alpha^2 \mathrm{V}(X)$, on déduit facilement les valeurs

$$\mathrm{E}(Y) = \frac{a+b}{2} \quad ; \quad \mathrm{V}(Y) = \frac{(b-a)(b-a+2)}{12}$$

Simulation informatique. `random(n)` retourne un nombre aléatoire suivant la loi uniforme sur {0, . . . , n − 1}. On en déduit la simulation d'une v.a de loi uniforme sur {a, . . . ,b} :

```
function unif(a,b :integer) :integer ;
begin unif:=a+random(b-a+1); end;
```

• Soit $X_1, \cdots, X_m$ des v.a suivant la loi uniforme sur $[\![1, n]\!]$, et indépendantes. Pour déterminer la loi du maximum M de ces variables, on commence par déterminer la probabilité $\mathrm{P}(M \leqslant k)$, pour $1 \leqslant k \leqslant m$:

$$\mathrm{P}(M \leqslant k) = \mathrm{P}\left(\bigcap_{i=1}^{m} (X_i \leqslant k)\right) = \prod_{i=1}^{m} \mathrm{P}(X_i \leqslant k) = \left(\frac{k}{n}\right)^m$$

en utilisant l'indépendance des X_i. De

$$(M \leqslant k - 1) \cup (M = k) = (M \leqslant k)$$

on déduit, par additivité (valable avec $k = 1$) :

$$\mathrm{P}(M = k) = \mathrm{P}(M \leqslant k) - \mathrm{P}(M \leqslant k - 1) = \left(\frac{k}{n}\right)^m - \left(\frac{k-1}{n}\right)^m$$



Variables aléatoires discrètes

8

1. Espaces probabilisés quelconques

Certaines expériences aléatoires conduisent à un espace probabilisable (ensemble des résultats de l'expérience) comprenant un nombre infini d'éléments. Par exemple, l'expérience qui consiste à lancer une pièce de monnaie jusqu'à l'obtention de deux « pile » consécutifs amène à un espace probabilisable infini, puisqu'il doit contenir au moins les événements :

$$P_1P_2\ ;\ F_1P_2P_3\ ;\ F_1F_2P_3P_4\ ;\ F_1F_2F_3P_4P_5\ \ ;\ \ldots.$$

(avec P_i (resp. F_i), les événements « pile (resp. face) au lancer n° i », et en omettant les signes d'intersection.

Cette situation oblige à redéfinir les notions d'espace probabilisé et de variable aléatoire.

1.1 Définitions

Tribu, espace probabilisable. Soit Ω un ensemble non vide quelconque. Une **tribu** sur Ω est un ensemble $\mathcal{T}$ de parties de Ω (c'est-à-dire un sous-ensemble de $\mathcal{P}(\Omega)$) tel que :

- $\Omega \in \mathcal{T}$;
- si $A \in \mathcal{T}$, alors $\overline{A} \in \mathcal{T}$;
- si, pour tout $n \in \mathbf{N}$, $A_n \in \mathcal{T}$, alors $\bigcup_{n \in N} A_n \in \mathcal{T}$.

Soit $\mathcal{T}$ une tribu sur Ω. $(\Omega, \mathcal{T})$ est un **espace probabilisable**.

On convient d'appeler **événement** tout élément de $\mathcal{T}$.

Probabilité. Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une **probabilité** sur Ω est une application $P : \mathcal{T} \to [0, 1]$, $A \mapsto P(A)$ telle que :

• $P(\Omega) = 1$;

• si les A_n sont deux à deux incompatibles, alors :

$$P\left(\bigcup_{n\in\mathbf{N}} A_n\right) = \sum_{n\in\mathbf{N}} P(A_n) \quad \text{(propriété de } \sigma - \textbf{additivité}\text{)}$$

Le triplet $(\Omega, \mathcal{T}, P)$ est alors un **espace probabilisé**.

Propriétés

• On retrouve les mêmes propriétés que pour les espaces probabilisés finis, voir § 7.2.2. Pour mémoire :

– $P(\varnothing) = 0$;

– $P\left(\overline{A}\right) = 1 - P(A)$;

– (additivité) si les A_i sont deux à deux incompatibles, alors

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

– (croissance) si $A \subset B$, alors

$$P(A) \leqslant P(B) \quad \text{et} \quad P\left(B \backslash A\right) = P(B) - P(A)$$

– formule du crible.

• On a de plus les propriétés suivantes, dites **propriétés de limite monotone :**

– si $A_0 \subset A_1 \subset \cdots \subset A_n \subset \cdots$, alors

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$$

– si $A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots$, alors

$$P\left(\bigcap_{n=0}^{+\infty} A_i\right) = \lim_{n\to+\infty} P(A_n)$$

• On lance une pièce de monnaie équilibrée indéfiniment. Soit A l'événement « on obtient indéfiniment "face" ». A est l'intersection des événements A_n : « les n premiers lancers donnent "face" », avec n appartenant à $\mathbf{N}^*$.

$$\forall n \in \mathbf{N}^*, \quad A_n \supset A_{n+1}.$$

Par indépendance des lancers, on a : $P(A_n) = \left(\frac{1}{2}\right)^n$

Par conséquent, par la deuxième propriété de limite monotone :

$$\mathrm{P}(A) = 0$$

• La probabilité p de « casser » un mot de passe de 6 caractères (quand on procède au hasard) est $p = 1/36^6$, soit $p \approx 4,6 \times 10^{-10}$. On fait des essais, soit B l'événement « le mot de passe est trouvé ».

B est la réunion des événements B_n : « le mot de passe est trouvé en au plus n essais », $n \in \mathbf{N}^*$.

$$\forall n \in \mathbf{N}^* \ , \quad B_n \subset B_{n+1}.$$

La probabilité de B_n est de $1-(1-p)^n$: en effet, l'événement contraitre de B_n est : « en n essais, le mot de passe n'a pas été trouvé ».

On conclut alors, en utilisant la première propriété de limite monotone, que $\mathrm{P}(B) = 1$.

On peut obtenir ces résultats dans le cadre de la loi géométrique, voir § 8.4.1.

• Soit $(X_n)_{n\in N}$ une suite quelconque d'événements. En appliquant la première propriété de limite monotone à la suite (A_n) définie par $A_n = \bigcup_{k=0}^{n} X_k$, on obtient

$$\mathrm{P}\left(\bigcup_{n=0}^{+\infty} X_n\right) = \lim_{n\to+\infty} \mathrm{P}\left(\bigcup_{k=0}^{n} X_k\right)$$

Et on a la formule analogue pour l'intersection.

• Un événement de probabilité 1 est dit **quasi-certain**. Un événement de probabilité 0 est dit **quasi-impossible**.

• On peut parler d'une suite (infinie) d'événements indépendants : une suite d'événements est indépendante ssi toute sous-suite finie extraite l'est. Les exemples fondamentaux : les résultats successifs d'une pièce de monnaie, d'un dé... ; les résultats successifs de tirages AVEC REMISE d'une boule dans une urne.

• La probabilité conditionnelle sachant B, avec $\mathrm{P}(B) \neq 0$, a même définition que dans le cas où Ω est fini :

$$\mathrm{P}_B(A) = \frac{\mathrm{P}(\mathrm{A} \cap \mathrm{B})}{\mathrm{P}(B)}$$

et même utilisation, la formule des probabilités composées :

$$\mathrm{P}(A \cap B) = \mathrm{P}_B(A)\,\mathrm{P}(B) \quad ;$$

$$\mathrm{P}(A_1 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\,\mathrm{P}_{A_1}(A_2)\cdots \mathrm{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n)$$

• La formule des probabilités totales se généralise avec des systèmes $(A_i)_{i\in I}$ (quasi-)complets d'événements, c'est-à-dire vérifiant:

Pour tout i, $\mathrm{P}(A_i) \neq 0$; si $i \neq j$, $A_i \cap A_j = \varnothing$; $\mathrm{P}\left(\bigcup_{i\in I} A_i\right) = 1$, avec I fini ou dénombrable.

On a, avec $(A_i)_{i\in I}$ système (quasi-) complet d'événements, pour tout événement B :

$$\mathrm{P}(B) = \sum_{i\in I} \mathrm{P}(B \cap A_i) = \sum_{i\in I} \mathrm{P}_{A_i}(B)\,\mathrm{P}(A_i)$$

Quand Ω (ensemble des résultats élémentaires de l'expérience aléatoire) est infini, on ne peut pas définir sans contradiction la probabilité $\mathrm{P}(A)$ pour tout élément A de $\mathcal{P}(\Omega)$. Il faut se restreindre à un sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$ convenablement structuré. Dans la pratique, cela ne pose pas de difficultés.

La propriété de σ–additivité demandée à la probabilité étend la propriété d'additivité à des réunions infinies. Elle comporte la somme d'une série. La convergence de cette série n'a pas à être établie : en effet les sommes partielles $\sum_{k=0}^{n} \mathrm{P}(A_k)$ forment une suite croissante (car $\mathrm{P}(A_k) \geqslant 0$ pour tout k), et majorée, car

$$\sum_{k=0}^{n} \mathrm{P}(A_k) = \mathrm{P}\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \mathrm{P}(\Omega) = 1$$

2. Variables aléatoires infinies discrètes

2.1 Définitions

Soit $(\Omega, \mathcal{T}, \mathrm{P})$ un espace probabilisé.

Variable aléatoire

• Une **variable aléatoire** (v.a) définie sur Ω est une application X de Ω dans $\mathbf{R}$ telle que, pour tout $x \in \mathbf{R}$, l'ensemble

$$(X \leqslant x) = \{\omega \;;\; \omega \in \Omega, X(\omega) \leqslant x\}$$

appartienne à $\mathcal{T}$.

• Une **variable aléatoire infinie discrète** (vaid) définie sur Ω est une v.a définie sur Ω telle que l'on ait :

$$X(\Omega) = \{x_i \;;\; i \in I\},$$

où I est l'ensemble $\mathbf{N}$, l'ensemble $\mathbf{Z}$, ou un sous-ensemble infini de $\mathbf{N}$ ou $\mathbf{Z}$.

• Une **variable aléatoire discrète** (vad) est une variable aléatoire finie ou infinie discrète.

Loi de probabilité

• Déterminer la loi de probabilité de la v.a infinie discrète X, c'est déterminer $X(\Omega)$, et, pour chaque $x_i \in X(\Omega)$, déterminer la probabilité $\mathrm{P}(X = x_i)$.

• Réciproquement, pour vérifier que les couples $(x_i, p_i)_{i \in I}$ définissent la loi de probabilité d'une vaid X, avec $\mathrm{P}(X = x_i) = p_i$, il suffit de vérifier :

– $\forall i \in I, p_i \geqslant 0$;

– la série de terme général $p_i, i \in I$, est convergente, de somme 1.

On dira aussi que X est bien une variable aléatoire.

2.2 Propriétés

Ce sont les mêmes que dans le cas fini :

La somme, le produit, le quotient (si le dénominateur ne s'annule pas) de deux vad définies sur le même ensemble Ω est une v.a définie sur Ω.

Si φ est une fonction à valeurs dans $\mathbf{R}$ définie sur $X(\Omega)$, la composée $\varphi \circ X$ est une v.a définie sur Ω.

L'ensemble des v.a définies sur Ω constitue un espace vectoriel pour les opérations $X + Y, \alpha X$.

2.3 Fonction de répartition

Définition. Soit X une v.a sur Ω. La **fonction de répartition** de la v.a X est l'application F_X de $\mathbf{R}$ dans $\mathbf{R}$ définie par :

$$\forall x \in \mathbf{R}\ ,\ F_X(x) = \mathrm{P}(X \leqslant x)$$

Propriétés.

• Soit F_X la f.r d'une v.a X. Alors

– F_X est croissante sur $\mathbf{R}$.

– F_X est continue à droite en tout point de $\mathbf{R}$.

– $\lim\limits_{x \to -\infty} F_X(x) = 0\ ;\ \lim\limits_{x \to +\infty} F_X(x) = 1$

• Réciproquement, toute application de $\mathbf{R}$ dans $\mathbf{R}$ vérifiant ces trois propriétés est la fonction de répartition d'une v.a X.

• Avec F_X la f. r de la v.a X, pour tout a, b appartenant à $\mathbf{R}$ tels que $a \leqslant b$:

$$\mathrm{P}(a < X \leqslant b) = F_X(b) - F_X(a)$$

• En particulier, si X est à valeurs dans $\mathbf{N}$, ou dans $\mathbf{Z}$, on a :

$$\mathrm{P}(X = k) = F_X(k) - F_X(k-1)$$

• Soit F_X la f.r d'une v.a X. La v.a X est infinie discrète (resp. finie) ssi F_X est une fonction en escalier (c'est-à-dire constante par intervalles), telle que l'ensemble de ses points de discontinuité est dénombrable (resp. fini). Cet ensemble est alors égal à $X(\Omega)$.

Indépendance de variables aléatoires discrètes

Définitions.

• Deux vad X, Y définies sur Ω sont dites **indépendantes** ssi :
$\forall x \in X(\Omega), \forall y \in Y(\Omega)$,

$$\mathrm{P}((X = x) \cap (Y = y)) = \mathrm{P}(X = x)\,\mathrm{P}(Y = y)$$

• Les vad $X_1, X_2, \cdots, X_n$ sont dites **mutuellement indépendantes** (ou **indépendantes**) ssi :
$\forall x_1 \in X_1(\Omega),\ \forall x_2 \in X_2(\Omega), \cdots, \forall x_n \in X_n(\Omega)$,

$$\mathrm{P}\left(\bigcap_{i=1}^{n}(X_i = x_i)\right) = \prod_{i=1}^{n}\mathrm{P}(X_i = x_i)$$

• La suite $(X_n)_{n\in\mathbf{N}^*}$ est une suite de v.a indépendantes ssi, pour tout $n \in \mathbf{N}^*$, les v.a $X_1, X_2, \cdots, X_n$ sont indépendantes.

Propriété

• Si $(X_n)_{n\in\mathbf{N}^*}$ est une suite de v.a indépendantes, alors toute fonction de $X_1, X_2, \cdots, X_n$ est indépendante de toute fonction de $X_{n+1}, \cdots, X_{n+p}$.

En particulier, si X et Y sont indépendantes, alors toute fonction de X est indépendante de toute fonction de Y.

2.4 Espérance (ou moyenne)

Définition. Soit X une variable aléatoire infinie discrète définie sur Ω.

On dit que X admet une **espérance** $\mathrm{E}(X)$ ssi la série de terme général $x_i P(X = x_i), x_i \in X(\Omega)$, est absolument convergente, et on pose alors :

$$\mathrm{E}(X) = \sum_{x_i \in X(\Omega)} x_i \mathrm{P}(X = x_i)$$

Un cas particulier d'utilisation fréquente :

$$\text{Si } X(\Omega) = \mathbf{N}, \quad \mathrm{E}(X) = \sum_{k=0}^{+\infty} k\mathrm{P}(X = k) ;$$

sous réserve de convergence absolue, c'est-à-dire ici de convergence, car la série est à termes positifs.

Propriétés

Linéarité de l'espérance

Si X et Y admettent des espérances, alors il en est de même pour $X + Y$ et $aX + b$ $(a, b \in \mathbf{R})$, et

$$E(X+Y) = E(X) + E(Y) ;$$

$$E(aX+b) = aE(X) + b.$$

L'espérance est donc une forme linéaire sur l'espace vectoriel des vad définies sur Ω et admettant une espérance.

Théorème de transfert

Soit X une vaid, φ une application définie sur $X(\Omega)$. Alors $\varphi(X)$ est une vaid, et, sous réserve de la convergence absolue de la série de terme général $\varphi(x_i)\,\mathrm{P}(X = x_i)$, $i \in X(\Omega)$:

$$\mathrm{E}(\varphi(X)) = \sum_{x_i \in X(\Omega)} \varphi(x_i)\,\mathrm{P}(X = x_i)$$

En particulier, pour $r \in \mathbf{N}^*$, et toujours sous réserve de convergence absolue :

$$\mathrm{E}(X^r) = \sum_{x_i \in X(\Omega)} x_i^r \mathrm{P}(X = x_i)$$

$\mathrm{m}_r(X) = \mathrm{E}(X^r)$ est appelé **moment d'ordre** r de X.

- Attention, certaines v.a infinies discrètes n'ont pas d'espérance.

Soit X une v.a dont la loi de probabilité est donnée par :

$$X(\Omega) = \mathbf{N}^* ; \ \forall n \in \mathbf{N}^*, \ \ \mathrm{P}(X = n) = \frac{6}{\pi^2 n^2}$$

On sait que la série de terme général $\frac{1}{n^2}$, $n \in \mathbf{N}^*$, est convergente (série de Riemann), et, étant admis que $\sum_{n=0}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, on établit sans peine que X est bien une variable aléatoire (ou que l'on a bien défini une loi

de probabilité). Mais cette v.a n'admet pas d'espérance. En effet la série de terme général $n\mathrm{P}(X = n)$ est divergente, car

$$n\mathrm{P}(X = n) = \frac{6}{\pi^2 n}$$

et la série de terme général $\frac{1}{n}$, $n \in \mathbf{N}^*$, est une série de Riemann divergente. Par conséquent, X n'admet pas d'espérance.

Variance, écart-type

Définitions

• Soit X une v.a infinie discrète admettant une espérance $\mathrm{E}(X)$. Sous réserve de convergence de la série, la **variance** de X est définie par :

$$\mathrm{V}(X) = \sum_{x_i \in X(\Omega)} [x_i - \mathrm{E}(X)]^2 \mathrm{P}(X = x_i)$$

D'après le théorème de transfert, on a alors

$$\mathrm{V}(X) = \mathrm{E}\left([X - \mathrm{E}(X)]^2\right)$$

• Pour une v.a X admettant une variance, l'**écart-type** $\sigma(X)$ est la racine carrée de la variance :

$$\sigma(X) = \sqrt{\mathrm{V}(X)}$$

Propriétés

• Pour toute v.a X admettant une variance, $\mathrm{V}(X) \geqslant 0$. Si $\mathrm{V}(X) = 0$, alors X est une v.a quasi-certaine : $X(\Omega) = \{c\}$; $\mathrm{P}(X = c) = 1$.

• Si X admet une variance, et $a, b \in \mathbf{R}$, alors $aX + b$ admet une variance, et

$$\mathrm{V}(aX + b) = a^2 \mathrm{V}(X)$$

Démonstration. Soit $Y = aX + b$. Y admet une espérance, et

$$Y - \mathrm{E}(Y) = a(X - \mathrm{E}(X))$$

X admet une variance $\mathrm{V}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right)$, ce qui donne le résultat.

Théorème de Koenig-Huygens

Soit X une v.a infinie discrète. X admet une variance ssi X admet un moment d'ordre 2, $\mathrm{E}\left(X^2\right)$, et on a alors :

$$\mathrm{V}(X) = \sum_{x_i \in X(\Omega)} {x_i}^2 \mathrm{P}(X = x_i) - [\mathrm{E}(X)]^2 = \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2$$

Démonstration. Montrons que si X admet un moment d'ordre 2, alors X admet une espérance ; on a $\left(|x_i| + 1\right)^2 \mathrm{P}(X = x_i) \geqslant 0$, donc, en développant le carré :

$$0 \leqslant |x_i| \mathrm{P}(X = x_i) \leqslant \frac{1}{2}\left(x_i^2 + 1\right) \mathrm{P}(X = x_i).$$

La série de terme général $\left(x_i^2 + 1\right) \mathrm{P}(X = x_i)$ est convergente comme somme de deux séries convergentes, donc la série de terme général $x_i \mathrm{P}(X = x_i)$ est absolument convergente, donc X admet une espérance. Le calcul suivant :

$$(X - [\mathrm{E}(X)])^2 = X^2 - 2X\,\mathrm{E}(X) + [\mathrm{E}(X)]^2$$

montre alors, en utilisant la linéarité de l'espérance, que X admet une variance ssi X admet un moment d'ordre 2, et prouve l'égalité à établir.

• Une v.a admettant une espérance peut ne pas admettre de moment d'ordre 2, et donc pas de variance.

> Soit X telle que
>
> $$\forall n \in \mathbf{N}^*, \mathrm{P}(X = n) = \frac{\mathrm{a}}{n^3}, \quad \text{avec} \quad \mathrm{a} = \sum_{n=1}^{+\infty} \frac{1}{n^3}$$
>
> X admet une espérance, car $n\mathrm{P}(X = n) = \frac{\mathrm{a}}{n^2}$, et la série de Riemann $\sum \frac{1}{n^2}$ est convergente ; mais X n'admet pas de moment d'ordre 2, car $n^2\mathrm{P}(X = n) = \frac{\mathrm{a}}{n}$, et la série de Riemann $\sum \frac{1}{n}$ est divergente. X n'admet donc pas de variance.

Variance d'une somme de 2 v.a infinies discrètes

Soit X et Y deux v.a infinies discrètes admettant des variances. Alors $X + Y$ admet une variance, et XY admet une espérance. De plus :

– si X et Y sont indépendantes, alors

$$\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y)$$

– dans le cas général :

$$\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y) + 2\,\mathrm{Cov}(X, Y),$$

avec

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y),$$

avec

$$\mathrm{E}(XY) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j \mathrm{P}\left(X = x_i, Y = y_j\right)$$

• **Généralisation à n v.a infinies discrètes**

Si les X_i admettent des variances, alors la somme $\sum_{i=1}^{n} X_i$ admet une variance, et :

– si les X_i sont deux à deux indépendantes, alors

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$$

– dans le cas général

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \operatorname{Cov}(X_i, Y_j)$$

3. Couple de variables aléatoires discrètes

3.1 Définitions

Soit X, Y deux variables aléatoires discrètes. Les définitions sont les mêmes que dans le cas fini, voir § 7. 3. 1. Pour mémoire :

• Déterminer la **loi du couple** (X, Y), c'est déterminer, pour chaque $x_i \in X(\Omega)$ et chaque $y_j \in Y(\Omega)$ la probabilité

$$P\left((X = x_i) \cap (y = y_j)\right)$$

• La loi du couple (X, Y) permet de trouver les **lois marginales** (les lois de X et de Y), en utilisant la formule de probabilités totales.

• Pour y_j fixé, la **loi conditionnelle** de la variable X sachant $(Y = y_j)$ est donnée par

$$P_{(Y=y_j)}(X = x_i) = \frac{P\left((X = x_i) \cap (Y = y_j)\right)}{P(Y = y_j)}$$

et de même pour la loi de la variable Y sachant $(X = x_i)$.

Théorème de transfert

Soit X, Y deux vaid, φ une fonction de deux variables. Sous réserve d'existence :

$$E(\varphi(X, Y)) = \sum_{(x_i, y_j) \in X(\Omega) \times Y(\Omega)} \varphi(x_i, y_j) P\left((X = x_i) \cap (Y = y_j)\right)$$

En particulier,

$$\mathrm{E}(XY) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j \mathrm{P}\left(X = x_i, Y = y_j\right)$$

sous réserve d'existence, c'est-à-dire sous réserve de convergence absolue de la série. Si X et Y admettent des variances, alors cette convergence est acquise (admis).

3.2 Covariance

Pour déterminer la variance de $X + Y$, on fait les mêmes calculs que dans le cas fini, mais sous réserve d'existence. On admet que ces calculs sont possibles si X et Y admettent des variances.

Définition. Soit X, Y deux vaid admettant des variances. La covariance de (X, Y) est le nombre réel

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

avec $\quad \mathrm{E}(XY) = \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j \mathrm{P}\left(X = x_i, Y = y_j\right)$

Propriétés. Ce sont les mêmes que dans le cas fini, sous réserve d'existence. Pour mémoire :

• Si X, Y sont deux vaid admettant des variances, alors $X + Y$ admet une variance, et

$$\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y) + 2\mathrm{Cov}(X, Y)$$

• $\mathrm{Cov}(X, Y)$ est de signe de quelconque.

• On a aussi $\mathrm{Cov}(X, Y) = \mathrm{E}([X - \mathrm{E}(X)]\,[Y - \mathrm{E}(Y)])$.

• Si X, Y, Z admettent des variances et si a est un nombre réel, on a :

– $\mathrm{Cov}(X, X) = \mathrm{V}(X)$;

– $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;

– $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$;

– $\mathrm{Cov}(aX, Y) = \mathrm{Cov}(X, aY) = a\mathrm{Cov}(X, Y)$.

• La formule $\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y) + 2\mathrm{Cov}(X, Y)$ permet le calcul de $\mathrm{Cov}(X, Y)$, si les variances de X, Y et $X + Y$ sont connues.

• Soient X, Y deux vaid admettant des variances, et indépendantes. On a alors les propriétés suivantes, équivalentes entre elles :

$$\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$$

$$\mathrm{Cov}(X, Y) = 0$$

$$V(X + Y) = V(X) + V(Y)$$

Les réciproques sont fausses. La covariance de deux vad non indépendantes peut être égale à 0. On peut juste dire que si la covariance de (X, Y) est différente de 0, alors X, Y ne sont pas indépendantes.

Coefficient de corrélation linéaire

Définition. Soit X, Y des vad admettant des variances, et non quasi-certaines. Le coefficient de corrélation linéaire est le nombre réel $r_{X,Y}$ défini par

$$r_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\,\sigma(Y)}$$

Propriétés

- $|r_{X,Y}| \leqslant 1$; si $|r_{X,Y}| = 1$, alors il existe a, b réels tels que l'événement $Y = aX + b$ est quasi-certain, et a est du signe de r.
- Si X, Y sont indépendantes, alors $r_{X,Y} = 0$. La réciproque est fausse.

4. Variables infinies discrètes usuelles

4.1 Loi géométrique

Définition. Soit $p \in \,]0\,;1[$. On dit que X suit la loi géométrique de paramètre p, et on note $X \hookrightarrow G(p)$, ssi :

$$X(\Omega) = \mathbf{N}^* \;\; ; \;\; \forall n \in \mathbf{N}^* \,, \quad P(X = n) = (1 - p)^{n-1} p$$

Modèle. On répète une épreuve indéfiniment. À chaque épreuve, la probabilité du succès est p fixé dans $]0\,;1[$. Les résultats des épreuves successives sont indépendants.

Soit X le temps d'attente du premier succès, c'est-à-dire le numéro de l'épreuve qui amène le premier succès. Alors

$$X \hookrightarrow \mathcal{G}(p)$$

Espérance, variance. Soit $X \hookrightarrow G(p)$. Alors

$$E(X) = \frac{1}{p} \quad ; \quad V(X) = \frac{1-p}{p^2}$$

- On retrouve facilement la loi géométrique à partir de son modèle. En notant E_i et S_i l'échec et le succès à l'épreuve n° i, on a, pour tout $n \in \mathbf{N}^*$:

$$(X = n) = E_1 \cap \cdots \cap E_{n-1} \cap S_n$$

d'où le résultat, par indépendance des résultats successifs.

• Vérifions qu'il s'agit d'une loi de probabilité. Avec $q = 1 - p$:

– $\forall n \in \mathbf{N}^*$, $\mathrm{P}(X = n) = q^{n-1}p \geqslant 0$;

– $\displaystyle\sum_{n=1}^{+\infty} P(X = n) = \sum_{n=1}^{+\infty} q^{n-1}p = p\sum_{n=0}^{+\infty} q^n = p\frac{1}{1-q} = 1$

car $q \in]0\,;1[$, la série géométrique $\displaystyle\sum_{n\in\mathbf{N}^*} q^n$ est donc convergente, de somme $\dfrac{1}{1-q}$. Voir § 2.4.2.

• La vérification faite que la loi géométrique est bien une loi de probabilité est tout à fait essentielle. Elle montre que dans une suite d'épreuves identiques et indépendantes, on est quasi-certain d'obtenir le succès.

• Les calculs de l'espérance et de la variance de X suivant la loi $\mathcal{G}(p)$ doivent être parfaitement maîtrisés. Il font appel aux calculs sur les séries géométriques et séries géométriques dérivées (§ 2.4.2). En voici les grandes lignes. On note $q = 1 - p$.

– Pour $\mathrm{E}(X)$: sous réserve de convergence (absolue) de la série :

$$\mathrm{E}(X) = \sum_{n=1}^{+\infty} n\mathrm{P}(X = n) = \sum_{n=1}^{+\infty} nq^{n-1}p = p\sum_{n=1}^{+\infty} nq^{n-1}$$

Or la série de terme général nq^{n-1} est convergente (série géométrique dérivée de raison $q \in]0\,;1[$), et a pour somme

$$\sum_{n=1}^{+\infty} nq^{n-1} = \frac{1}{(1-q)^2} = \frac{1}{p^2}$$

D'où la conclusion. Ce résultat est facile à retenir, il signifie par exemple que quand on lance un dé cubique équilibré, le « 6 » (probabilité d'apparition égale à $\frac{1}{6}$) apparaît en moyenne tous les 6 lancers.

– Pour la variance, on commence par écrire (sous réserve de convergence absolue des séries rencontrées) :

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2 = \mathrm{E}(X(X-1) + X) - [\mathrm{E}(X)]^2\\ &= \mathrm{E}(X(X-1)) + \mathrm{E}(X) - [\mathrm{E}(X)]^2 \text{ (linéarité de l'éspérance)}\end{aligned}$$

On utilise le théorème de transfert, toujours sous réserve de convergence (absolue) :

$$\mathrm{E}(X(X-1)) = \sum_{n=1}^{+\infty} n(n-1)\,q^{n-1}p = pq\sum_{n=2}^{+\infty} n(n-1)\,q^{n-2}$$

La série obtenue est convergente (série géométrique dérivée seconde de raison $q \in]0\,;1[$), de somme $\frac{2}{(1-q)^3} = \frac{2}{p^3}$. Il ne reste plus qu'à achever les calculs.

• Pour une simulation informatique d'une v.a suivant une loi géométrique, voir § 10.2.3.

• **Variantes de la loi géométrique.** On considère toujours une succession d'épreuves identiques et indépendantes, la probabilité du succès à chaque épreuve étant p. Les lois des v.a suivantes ne sont pas au programme, mais elles inspirent de nombreux exercices.

– **Nombre d'échecs avant le premier succès.** Soit X ce nombre. On a $X(\Omega) = \mathbf{N}$, et, pour tout $n \in \mathbf{N}$,

$$(X = n) = E_1 \cap \cdots \cap E_n \cap S_{n+1},$$

donc, par indépendance : $\forall n \in \mathbf{N}, \mathrm{P}(X = n) = q^n p.$

Il s'agit d'une loi de probabilité car pour tout $n, q^n p \geqslant 0$, et

$$\sum_{n=0}^{+\infty} q^n p = p\frac{1}{1-q} = 1.$$

Pour le calcul de l'espérance et de la variance, on peut procéder directement, on peut aussi remarquer que la va $Y = X + 1$ suit la loi géométrique de paramètre p : $Y(\Omega) = \mathbf{N}^*$, et, pour tout n dans $\mathbf{N}^*$: $\mathrm{P}(Y = n) = \mathrm{P}(X = n-1) = q^{n-1}p$. On utilise alors les formules : $\mathrm{E}(aX+b) = a\mathrm{E}(X) + b, \mathrm{V}(aX+b) = a^2\mathrm{V}(X)$ pour trouver :

$$\mathrm{E}(X) = \frac{q}{p}\ ; \quad \mathrm{V}(X) = \frac{q}{p^2}$$

Mentionnons qu'il s'agit d'une loi sans mémoire :

$$\forall n, m \in \mathbf{N}\ ,\ \mathrm{P}_{(X \geq n)}(X \geq n + m) = \mathrm{P}(X \geq m)\,.$$

– **Temps d'attente du r-ème succès ($r \in \mathbf{N}^*$).** Soit X ce nombre. X est à valeurs dans l'ensemble $\{r\,; r+1\,;\ \cdots\}$, et pour tout $n \geqslant r$, $(X = n)$ est l'intersection des deux événements « $r - 1$ succès au cours des $n - 1$ premières épreuves », et « succès à la n-ème épreuve ». Par indépendance des épreuves, on a donc :

$$\mathrm{P}(X = n) = \binom{n-1}{r-1} q^{n-1-(r-1)}p^r \times p = \binom{n-1}{r-1} q^{n-r}p^r$$

– **Loi géométrique tronquée.** X est le numéro de l'épreuve qui apporte le premier succès, mais le nombre d'épreuves est limité à n, n fixé. Si aucune épreuve n'apporte le succès, on convient que $X = 0$.

Pour $k \in \{1, \cdots, n\}$, on a $\mathrm{P}(X = k) = q^{k-1}p$, et $\mathrm{P}(X = 0) = q^n$. Il s'agit bien d'une loi de probabilité, car d'après l'identité géométrique :

$$\sum_{k=0}^{n} \mathrm{P}(X = n) = q^n + \sum_{k=1}^{n} q^{k-1}p = q^n + p\frac{1 - q^n}{1 - q} = 1$$

Bien remarquer qu'il faut traiter à part le cas $(X = 0)$.

Pour le calcul de l'espérance, la démarche est à connaître :

$$\mathrm{E}(X) = \sum_{k=1}^{n} kq^{k-1}p = pf'(p)\,, \text{ avec } \quad f(p) = \sum_{k=1}^{n} q^k = q\frac{1 - q^n}{1 - q}$$

On calcule alors la dérivée de f en utilisant la deuxième expression.

4.2 Loi de Poisson

Pas de modèle d'urne pour la loi de Poisson, qui est une loi-limite : quand n est « grand », et p « petit », une v.a X suivant la loi binomiale de paramètres n, p suit approximativement la loi de Poisson de paramètre np.

Définition. Soit $\lambda > 0$. On dit que X suit la **loi de Poisson de paramètre** λ, et on note $X \hookrightarrow \mathcal{P}(\lambda)$, ssi

$$X(\Omega) = \mathbf{N}\,;\, \forall n \in \mathbf{N}, \mathrm{P}(X = n) = \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$$

Espérance, variance. Soit $X \hookrightarrow \mathcal{P}(\lambda)$. Alors

$$\mathrm{E}(X) = \lambda\,;\, \mathrm{V}(X) = \lambda$$

Stabilité. Soit $X \hookrightarrow \mathcal{P}(\lambda)$, $Y \hookrightarrow \mathcal{P}(\mu)$, $\lambda > 0, \mu > 0$, X et Y indépendantes. Alors

$$X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$$

• Vérifions que la loi de Poisson $\mathcal{P}(\lambda)$ est une loi de probabilité. $\mathrm{P}(X = n)$ est toujours positif, et

$$\sum_{n=0}^{+\infty} P(X = n) = \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1$$

• Pour le calcul de l'espérance, on a, sous réserve de convergence (absolue) :

$$\mathrm{E}(X) = \sum_{n=0}^{+\infty} n\mathrm{P}(X = n) = \sum_{n=1}^{+\infty} n\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda}\sum_{n=1}^{+\infty}\lambda\frac{\lambda^{n-1}}{(n-1)!}$$

$$= \mathrm{e}^{-\lambda}\lambda\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda}\lambda e^{\lambda} = \lambda$$

Chaque étape du calcul est importante à comprendre : définition ; la somme démarre vraiment avec $n = 1$; simplification par n ; mise en facteur de λ et changement d'indice $k = n - 1$; on reconnaît la série exponentielle et les calculs précédents sont justifiés.

• Pour le calcul de $\mathrm{V}(X)$, on commence par procéder comme pour la loi géométrique :

$$\mathrm{V}(X) = \mathrm{E}(X(X-1)) + \mathrm{E}(X) - [\mathrm{E}(X)]^2$$

On utilise le théorème de transfert pour calculer $\mathrm{E}(X(X-1))$; la somme démarre valablement à $n = 2$; simplification par $n(n-1)$; changement d'indice $k = n - 2$.

• Établissons le résultat sur la stabilité : si $X \hookrightarrow \mathcal{P}(\lambda)$, $Y \hookrightarrow \mathcal{P}(\mu)$, X et Y indépendantes, alors $Z = X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$. Z prend ses valeurs dans $\mathbf{N}$, et, pour tout $n \in \mathbf{N}$:

$$(Z = n) = \bigcup_{k=0}^{n}((X = k) \cap (Y = n - k))$$

Par additivité, puis indépendance, il vient

$$\mathrm{P}(Z = n) = \sum_{k=0}^{n}\mathrm{P}(x = k)\,\mathrm{P}(Y = n - k) = \sum_{k=0}^{n}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}\mathrm{e}^{-\mu}\frac{\mu^{n-k}}{(n-k)!}$$

$$\mathrm{P}(Z = n) = \mathrm{e}^{-(\lambda+\mu)}\frac{1}{n!}\sum_{k=0}^{n} n!\frac{\lambda^k}{k!}\frac{\mu^{n-k}}{(n-k)!} = \mathrm{e}^{-(\lambda+\mu)}\frac{1}{n!}(\lambda + \mu)^n$$

d'après la formule du binôme. Et on a la conclusion.

• On généralise par récurrence la stabilité de la loi de Poisson à n v.a X_k indépendantes et suivant les lois $\mathcal{P}(\lambda_k)$: c'est acquis pour $n = 2$, et supposant que c'est acquis pour n, alors c'est acquis pour $n+1$, car $\sum_{k=1}^{n} X_k$

suit la loi de Poisson de paramètre $\sum_{k=1}^{n} \lambda_k$, et est indépendante de X_{n+1}, et on a la conclusion.

• Aspect numérique

Soit $X \hookrightarrow \mathcal{P}(\lambda)$, $\lambda > 0$. On vérifie facilement que

$$\forall n \in \mathbf{N}, \mathrm{P}(X = n+1) = \frac{\lambda}{n+1} \mathrm{P}(X = n)$$

En effet,

$$\forall n \in \mathbf{N}, \frac{\lambda}{n+1} \mathrm{P}(X = n) = \frac{\lambda}{n+1} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda} \frac{\lambda^{n+1}}{(n+1)!}$$

Grâce à cette formule, on peut écrire un programme en langage PASCAL affichant les tables numériques de la loi de Poisson pour une valeur donnée par l'utilisateur du paramètre λ.

La valeur de λ donné par l'utilisateur est stockée dans la variable d'entrée `lambda`. Avec $X \hookrightarrow \mathcal{P}(\lambda)$, et F la f.r de X, le programme affiche les valeurs successives de n, $\mathrm{P}(X = n)$, $F(n)$, tant que $F(n) < 1 - 10^{-6}$. Les variables de sortie sont `n, p, F`.

```
program poisson ;
var n : integer ; p,F:real ;
BEGIN
write('lambda=') ;readln(lambda) ;
n:=0 ;p:=exp(-lambda) ;F:=exp(-lambda) ;
writeln(n,p,F) ; {P(X = 0) = F(0) = e^{-λ}}
while F<1-1.e-6 do
begin
n:=n+1 ;
p:=p*lambda/n ; {utilise P(X = n+1) = λ/(n+1) P(X = n)}
 F:=F+p ; {utilise F(n+1) = F(n) + P(X = n+1)}
writeln(n,p,F) ;
end ;
END.
```

• Exemple de loi conditionnelle

Soit $X \hookrightarrow \mathcal{P}(\lambda)$, $Y \hookrightarrow \mathcal{P}(\mu)$, X et Y indépendantes. On sait que $Z = X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$. Montrons que la loi de X conditionnée par $(Z = n)$, n fixé dans $\mathbf{N}^*$, est une loi binomiale :

Pour tout $k \in \{0, \cdots, n\}$, on a :

$$\begin{aligned}
\mathrm{P}_{(\mathrm{Z=n})}(X=k) &= \frac{\mathrm{P}\left[(X=k)\cap(Z=n)\right]}{\mathrm{P}(Z=n)} = \frac{\mathrm{P}\left[(X=k)\cap(Y=n-k)\right]}{\mathrm{P}(Z=n)} \\
&= \frac{\mathrm{P}(X=k)\times\mathrm{P}(Y=n-k)}{\mathrm{P}(Z=n)} = \frac{\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}\mathrm{e}^{-\mu}\dfrac{\mu^{n-k}}{(n-k)!}}{\mathrm{e}^{-(\lambda+\mu)}\dfrac{(\lambda+\mu)^n}{n!}} \\
&= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^{\mathrm{k}}\left(\frac{\mu}{\lambda+\mu}\right)^{n-k}
\end{aligned}$$

D'où la conclusion. On est passé de l'événement $(X = k) \cap (Z = n)$ à l'événement $(X = k) \cap (Y = n - k)$ (ces deux événements sont bien égaux, car $Z = X + Y$), pour utiliser l'indépendance de X, Y (alors que X et Z ne sont pas indépendantes).

• Exemple d'utilisation de la formule des probabilités totales

Soit $N \hookrightarrow \mathcal{P}(\lambda)$ et Y une v.a telle que, pour tout $n \in \mathbf{N}^*$, $Y/(N = n)$ suive la loi $\mathcal{B}(n, p)$. En d'autres termes, on a, pour $0 \leq k \leq n$:

$$\mathrm{P}_{(N=n)}(X = k) = \binom{n}{k} p^k q^{n-k}$$

Alors Y suit la loi de Poisson de paramètre λp. On le montre en utilisant la formule des probabilités totales, avec le sce $(N = n)_{n\in\mathbf{N}}$:

$$\begin{aligned}
\mathrm{P}(X=k) &= \sum_{n=0}^{+\infty}\mathrm{P}_{(N=n)}(X=k)\,\mathrm{P}(N=n) \\
&= \sum_{n=k}^{+\infty}\binom{n}{k}p^kq^{n-k}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = p^k\mathrm{e}^{-\lambda}\frac{1}{k!}\sum_{n=k}^{+\infty}q^{n-k}\frac{\lambda^n}{(n-k)!} \\
&= p^k\mathrm{e}^{-\lambda}\frac{1}{k!}\sum_{i=0}^{+\infty}q^i\frac{\lambda^{i+k}}{i!} = \lambda^kp^k\mathrm{e}^{-\lambda}\frac{1}{k!}\sum_{i=0}^{+\infty}\frac{(q\lambda)^i}{i!} \\
&= (\lambda p)^k\mathrm{e}^{-\lambda}\frac{1}{k!}\mathrm{e}^{q\lambda} = \mathrm{e}^{-\lambda p}\frac{(\lambda p)^k}{k!}
\end{aligned}$$

Ce calcul concentre bien des difficultés techniques, il est donc à bien comprendre et maîtriser : formule des probabilités totales ; la somme démarre valablement à $n = k$; on met en facteur tout ce qui ne dépend pas de n et on simplifie les factorielles ; changement d'indice $i = n - k$; mise en facteur de λ^k ; série exponentielle ; règles de calculs sur les puissances.

• **Exemple de convergence en loi**

Soit $X_n \hookrightarrow \mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$, $n \in \mathbf{N}^*$, $\lambda > 0$. Alors, pour k fixé dans $\mathbf{N}$:

$$\lim_{n \to +\infty} \mathrm{P}(X_n = k) = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}$$

En effet, on a

$$\begin{aligned} \mathrm{P}(X_n = k) &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n!}{k!\,(n-k)!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

Or, d'une part :

$$\frac{n!}{k!\,(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!} \underset{n \to +\infty}{\sim} \frac{n^k}{k!}$$

Et d'autre part :

$$\left(1 - \frac{\lambda}{n}\right)^{n-k} = \mathrm{e}^{(n-k)\ln\left(1 - \frac{\lambda}{n}\right)} \xrightarrow[n \to +\infty]{} \mathrm{e}^{-\lambda}$$

car l'exposant est équivalent à $n \times \left(-\frac{\lambda}{n}\right)$, sa limite est donc $-\lambda$.

Et on a la conclusion.

On dit que la suite (X_n) **converge en loi** vers une v.a X de loi $\mathcal{P}(\lambda)$. (Voir § 9.3.3.)

Pratiquement, cela veut dire que pour les grandes valeurs de n, $\mathrm{P}(X = k)$ est une bonne approximation de $\mathrm{P}(X_n = k)$.



Variables aléatoires à densité Convergences, approximations estimation

9

1. Généralités

1.1 Définitions

Densité de probabilité

Une **densité de probabilité** est une fonction $f : \mathbf{R} \to \mathbf{R}$ telle que :

- $f \geqslant 0$ sur $\mathbf{R}$;
- f est continue sur $\mathbf{R}$ privé d'un nombre fini de points ;
- $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$

Soit f une densité de probabilité. La v.a aléatoire X est une **variable aléatoire à densité**, ou **variable aléatoire absolument continue**, (vaac) admettant f pour densité ssi :

– $X(\Omega) \subset \mathbf{R}$;

– pour tout a, b tels que $-\infty \leqslant a \leqslant b < +\infty$,

$$P(a \leqslant X \leqslant b) = \int_a^b f(t)\,\mathrm{d}t$$

Fonction de répartition

Comme pour n'importe quelle v.a, la **fonction de répartition** (f.r) de X est la fonction F_X définie par :

$$\forall x \in \mathbf{R}, \quad F_X(x) = \mathrm{P}(X \leqslant x)$$

Indépendance

On dit que les vaac X et Y sont **indépendantes** ssi

$$\forall x, y \in \mathbf{R}, \quad \mathrm{P}((X \leqslant x) \cap (Y \leqslant y)) = \mathrm{P}(X \leqslant x)\,\mathrm{P}(Y \leqslant y),$$

ou, de façon équivalente :

$$\forall x, y \in \mathbf{R}, \quad \mathrm{P}((X > x) \cap (Y > y)) = \mathrm{P}(X > x)\,\mathrm{P}(Y > y).$$

1.2 Propriétés, calculs

Premières propriétés

Soit X une vaac, et soit a, b dans $\mathbf{R}$ tels que $a \leqslant b$. On a :

- $\mathrm{P}(X = a) = 0$;
- $\mathrm{P}(a \leqslant X \leqslant b) = \mathrm{P}(a < X \leqslant b) = \mathrm{P}(a \leqslant X < b) = \mathrm{P}(a < X < b)$.

Probabilités, densité, fonction de répartition

Proposition 1. Soit X de densité f, de f.r F_X.
Alors, pour tout $x \in \mathbf{R}$:

$$\mathrm{P}(X \leqslant x) = F_x(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

On peut donc utiliser comme suit la fonction de répartition pour calculer des probabilités :
Pour tout $a, b \in \mathbf{R}$ tels que $a \leqslant b$,

$$\mathrm{P}(a \leqslant X \leqslant b) = \int_{a}^{b} f(t)\,\mathrm{d}t = F_X(b) - F_X(a) \;;$$

$$\mathrm{P}(X \leqslant b) = \int_{-\infty}^{b} f(t)\,\mathrm{d}t = F_X(b) \;;$$

$$\mathrm{P}(X > a) = \int_{a}^{+\infty} f(t)\,\mathrm{d}t = 1 - P(X \leqslant a) = 1 - F_X(a) \;.$$

Propriétés de la fonction de répartition d'une vaac

Proposition 2. Soit X une v.a de densité f, de f.r F_X.

- F_X est croissante ;
- $\lim\limits_{-\infty} F_X = 0$; $\lim\limits_{+\infty} F_X = 1$;
- F_X est continue sur tout $\mathbf{R}$;

- En tout point de $\mathbf{R}$ où f est continue (c'est-à-dire sur tout $\mathbf{R}$ sauf en un nombre fini de points), F_X est de classe C^1, et $F'_X = f$.
- Réciproquement, soit F une fonction croissante et continue sur $\mathbf{R}$, de classe C^1 sur $\mathbf{R}$ privé d'un nombre fini de points, et telle que

$$\lim_{-\infty} F_X = 0, \quad \lim_{+\infty} F_X = 1.$$

Alors F est la fonction de répartition d'une v.a X à densité, de densité $f = F'$ partout où F est de classe C^1.

Vous utiliserez la proposition 1 pour déterminer la fonction de répartition de X quand la densité de X est donnée.

Pour montrer qu'une fonction F donnée est la f.r d'une vaac X, vous devrez vérifier tous les points de la partie réciproque de la proposition 2. On trouve alors une densité en calculant $F'(x)$ partout où F est dérivable (c'est-à-dire sur $\mathbf{R}$ sauf en un nombre fini de points) et en prolongeant F' de façon arbitraire en ces points.

Si on sait que F est la fonction de répartition d'une v.a X et qu'on veut montrer que X est une vaac, il suffit de vérifier que :

- F est continue sur tout $\mathbf{R}$;
- F est de classe C^1 sur $\mathbf{R}$ privé d'un nombre fini de points.

Soit X et Y deux vaac indépendantes, de même densité f définie par :

$$f(t) = 0 \text{ si } t < 0 \ ; \quad f(t) = e^{-t} \text{ si } t \geqslant 0.$$

Soit M la v.a définie sur $\mathbf{R}$ par : $(M \leqslant x) = (X \leqslant x) \cap (Y \leqslant y)$. La variable aléatoire ainsi définie est le **maximum** de X, Y.

– f est une densité de probabilité, car f est $\geqslant 0$, continue sur $]-\infty, 0[$ et $]0\,, +\infty[$. f est nulle en dehors de $[0\,;\, +\infty[$, donc

$$\begin{aligned}\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t &= \int_0^{+\infty} e^{-t}\,\mathrm{d}t \\ &= \lim_{x\to+\infty} \int_0^x e^{-t}\,\mathrm{d}t \\ &= \lim_{x\to+\infty} \left[-e^{-t}\right]_0^x = 1\end{aligned}$$

– Déterminons la fonction de répartition de X et Y : f est nulle en dehors de $[0\,;\, +\infty[$, donc $F(x) = 0$ si $x < 0$;

Si $x \geqslant 0$, $F(x) = \int_0^x e^{-t}\,\mathrm{d}t = 1 - e^{-x}$.

– Pour montrer que M est une variable aléatoire à densité, on détermine d'abord la f.r F_M de M, puis on montre que M est une vaac, puis on détermine une densité f_M de M. Pour tout x appartenant à $\mathbf{R}$,

$$\begin{aligned} F_M(x) &= \mathrm{P}(M \leqslant x) \\ &= \mathrm{P}[(X \leqslant x) \cap (Y \leqslant x)] \\ &= \mathrm{P}(X \leqslant x)\,\mathrm{P}(Y \leqslant x), \end{aligned}$$

car X et Y sont indépendantes. On obtient donc :

$$F_M(x) = [F(x)]^2 = \begin{cases} 0 & \text{si } x < 0 \\ \left(1 - e^{-x}\right)^2 & \text{si } x \geqslant 0 \end{cases}$$

F_M est continue sur tout $\mathbf{R}$, de classe C^1 sur $]-\infty, 0[$ et $]0, +\infty[$, donc M est une variable aléatoire à densité. Pour avoir une densité f_M de M, on dérive F_M sur chacun des intervalles $]\infty\,; 0[$, $]0\,;\ +\infty[$ et on prolonge de manière arbitraire en 0. On a donc par exemple :

$$f_M(x) = \begin{cases} 0 & \text{si } x < 0 \\ 2e^{-x}\left(1 - e^{-x}\right) & \text{si } x \geqslant 0 \end{cases}$$

1.3 Espérance, variance

Espérance (ou moyenne)

Soit X de densité f. Par définition,

$$E(X) = \int_{-\infty}^{+\infty} t f(t)\,\mathrm{d}t$$

sous réserve de convergence de l'intégrale.

Théorème de transfert. Soit X de densité f, soit φ une fonction définie sur $\mathbf{R}$, continue sauf en un nombre fini de points. Alors, sous réserve de convergence absolue de l'intégrale :

$$E(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(t) f(t)\,\mathrm{d}t$$

Linéarité de l'espérance. Soit X et Y des v.a quelconques (discrètes ou à densité), admettant des espérances, et a, b deux réels. Alors $X + Y$ et $aX + b$ admettent des espérances, et on a

$$\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y) \quad ; \quad \mathrm{E}(aX + b) = aE(X) + b$$

Variance

Soit X de densité f, admettant une espérance $E(X)$. Par définition,

$$V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)\,\mathrm{d}t = \mathrm{E}\left([X - \mathrm{E}(X)]^2\right)$$

sous réserve de convergence de l'intégrale.

Soit X une va discrète ou à densité, admettant une variance, et a, b deux réels. Alors $aX + b$ admet une variance, et

$$\mathrm{V}(aX + b) = a^2\mathrm{V}(X)$$

Théorème de Koenig-Huygens. Soit X de densité f. X admet une espérance et une variance ssi X admet un moment d'ordre 2, c'est-à-dire ssi l'intégrale

$$m_2(X) = \mathrm{E}\left(X^2\right) = \int_{-\infty}^{+\infty} t^2 f(t)\,\mathrm{d}t$$

est convergente. On a alors $\quad \mathrm{V}(X) = \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2$

2. Variables aléatoires à densité usuelles

2.1 Loi uniforme

Elle modélise les situations telles que : on tire un nombre au hasard entre 0 et 1 . . .

Soit $a, b \in \mathbf{R}$ tels que $a < b$. On dit que X suit la **loi uniforme** sur $[a, b]$, et on note $X \hookrightarrow \mathcal{U}([a\,;\,b])$ ssi X admet pour densité la fonction f telle que :

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } a \leqslant x \leqslant b \\ 0 & \text{sinon} \end{cases}$$

Fonction de répartition

$$F(x) = \begin{cases} 0 & \text{si } x \leqslant a \\ \dfrac{x-a}{b-a} & \text{si } a \leqslant x \leqslant b \\ 1 & \text{si } x \geqslant b \end{cases}$$

Espérance

$$\mathrm{E}(X) = \frac{a+b}{2}$$

- Si $X \hookrightarrow \mathcal{U}([0\,;\,1])$, alors X pour densité f telle que
$$f(x) = 1 \text{ si } x \in [0, 1]\ ; \quad f(x) = 0 \text{ sinon}$$
et pour fonction de répartition F_X telle que
$$F_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ x & \text{si } 0 \leqslant x \leqslant 1 \\ 1 & \text{si } x \geqslant 1 \end{cases}$$
Une telle variable est simulée par la fonction « `random` » du langage PASCAL.
- Si $X \hookrightarrow \mathcal{U}([0\,;\,1])$, alors $Y = a + (b-a)X \hookrightarrow \mathcal{U}([a, b])$ $(a < b)$. Pour établir ce résultat, on détermine F_Y, f.r de Y. Pour tout $x \in \mathbf{R}$:
$$F_Y(x) = \mathrm{P}(Y \leqslant x) = \mathrm{P}(a + (b-a)X \leqslant x) = \mathrm{P}\left(X \leqslant \frac{x-a}{b-a}\right),$$
donc
$$F_Y(x) = F_X\left(\frac{x-a}{b-a}\right) = \begin{cases} 0 & \text{si } x \leqslant a \\ \dfrac{x-a}{b-a} & \text{si } a \leqslant x \leqslant b \\ 1 & \text{si } x \geqslant b \end{cases}$$
Car $\varphi : x \mapsto \frac{x-a}{b-a}$ est une bijection strictement croissante de $\mathbf{R}$ sur $\mathbf{R}$, avec $\varphi(a) = 0$ et $\varphi(b) = 1$. On a bien $Y \hookrightarrow \mathcal{U}([a, b])$.

Si $X \hookrightarrow \mathcal{U}([0\,;\,1])$, alors $Y = a + (b-a)X \hookrightarrow \mathcal{U}([a, b])$ $(a < b)$.
On utilise ce résultat pour simuler une v.a suivant la loi $\mathcal{U}([a, b])$.

```
function unif(a,b :real) :real ;
begin unif:=a+(b-a)*random;end;
```

2.2 Loi exponentielle

Elle modélise la durée de fonctionnement normal de certains appareils.

Soit $\alpha > 0$. On dit que X suit **la loi exponentielle** de paramètre α, et on note $X \hookrightarrow \mathcal{E}(\alpha)$, ssi X admet pour densité f telle que :
$$f(t) = \begin{cases} 0 & \text{si } t < 0 \\ \alpha e^{-\alpha t} & \text{si } t \geqslant 0 \end{cases}$$

Fonction de répartition
$$F(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - e^{-\alpha x} & \text{si } x \geqslant 0 \end{cases}$$

Espérance, variance

$$\mathrm{E}(X) = \frac{1}{\alpha}\,; \quad \mathrm{V}(X) = \frac{1}{\alpha^2}$$

• **Propriété caractéristique de la loi exponentielle.** Soit X une v.a suivant une loi exponentielle. On vérifie sans peine que

$$(1) \quad \forall t > 0, \forall t' > 0, \quad \mathrm{P}_{(X>t)}\left(X > t + t'\right) = \mathrm{P}\left(X > t'\right)$$

Réciproquement, on admet que, si X une v.a à densité prenant ses valeurs dans $[0\ ;\ +\infty[$, et vérifiant (1), alors X suit une loi exponentielle. On dit que la loi exponentielle est sans mémoire.

• Soit X de loi uniforme sur $]0\,;\,1]$, $\alpha > 0$, et Y la v.a définie par

$$Y = -\frac{\ln(X)}{\alpha}$$

Montrons que Y est une vaac de loi $\mathcal{E}(\alpha)$, en déterminant sa f.r G. Pour tout $x \in \mathbf{R}$,

$G(x) = \mathrm{P}(Y \leqslant x) = \mathrm{P}(-\ln(X) \leqslant \alpha x)$.

Avec F f.r de X, on a :

$$G(x) = \mathrm{P}\left(X \geqslant \mathrm{e}^{-\alpha x}\right) = 1 - F\left(\mathrm{e}^{-\alpha x}\right)$$
$$= \begin{cases} 1 - 0 & \text{si } \mathrm{e}^{-ax} \leqslant 0 \\ 1 - \mathrm{e}^{-ax} & \text{si } 0 \leqslant \mathrm{e}^{-ax} \leqslant 1 \\ 1 - 1 & \text{si } \mathrm{e}^{-ax} \geqslant 1 \end{cases}$$

Or pour tout $x \in \mathbf{R}$, $\mathrm{e}^{-\alpha x} > 0$, et $\mathrm{e}^{-\alpha x} \leqslant 1 \Leftrightarrow x \geqslant 0$.

Par conséquent : $G(x) = 1 - \mathrm{e}^{-\alpha x}$ si $x \geqslant 0$; $G(x) = 0$ si $x \leqslant 0$.

On reconnaît bien l'expression de la f.r d'une v.a de loi $\mathcal{E}(\alpha)$.

Si $X \hookrightarrow \mathcal{U}(]0\,;\,1])$, alors $Y = -\frac{\ln(X)}{\alpha} \hookrightarrow \mathcal{E}(\alpha)$ $(\alpha > 0)$.

On utilise ce résultat pour simuler une v.a suivant la loi $\mathcal{E}(\alpha)$:

```
 rogram simulex  ;
var al ha,U,Y:real;
BEGIN readln(al ha);
re eat U:= random until (U>0);
Y:=-ln(U) al ha; write(Y);END.
```

• Déterminons l'expression de la f.r F à partir de celle de la densité f :
Si $x < 0$, alors

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{x} 0\,\mathrm{d}t = 0\,;$$

Si $x \geqslant 0$, alors

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{0} 0\,\mathrm{d}t + \int_{0}^{x} \alpha e^{-\alpha t}\,\mathrm{d}t \\ &= 0 + \left[-e^{-\alpha t}\right]_{0}^{x} = 1 - e^{-\alpha x} \end{aligned}$$

Et on vérifie que la fonction obtenue est continue sur $\mathbf{R}$.
• Calculons $\mathrm{E}(X)$. Pour $x > 0$, on effectue l'intégration par parties :

$$\int_{0}^{x} t\alpha e^{-\alpha t}\,\mathrm{d}t = \left[-te^{-\alpha t}\right]_{0}^{x} + \int_{0}^{x} e^{-\alpha t}\,\mathrm{d}t = -xe^{-\alpha x} + \left[-\frac{1}{\alpha}e^{-\alpha t}\right]_{0}^{x}$$

On obtient donc, car f est nulle en dehors de $[0\,;\,+\infty[$:

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} tf(t)\,\mathrm{d}t = \int_{-0}^{+\infty} tf(t)\,\mathrm{d}t = \lim_{x\longrightarrow+\infty} \int_{0}^{x} t\alpha e^{-\alpha t}\,\mathrm{d}t = \frac{1}{\alpha}$$

• On procède de même pour le calcul de $\mathrm{E}\left(X^2\right)$, pour obtenir $\mathrm{V}(X)$.

Remarques

• Loi exponentielle et loi géométrique. Soit $X \hookrightarrow \mathcal{E}(\alpha)$, et T la v.a définie sur $\mathbf{N}^*$ par :

$$\forall n \in \mathbf{N}^*, (T = n) = (n - 1 \leqslant X \leqslant n)$$

Alors T suit la loi géométrique de paramètre $1 - \mathrm{e}^{-\alpha}$, car

$$\begin{aligned} \forall n \in \mathbf{N}^*, \mathrm{P}(T = n) &= F_X(n) - F_X(n-1) = -\mathrm{e}^{-\alpha n} + \mathrm{e}^{-\alpha(n-1)} \\ &= \mathrm{e}^{-\alpha(n-1)}\left(1 - \mathrm{e}^{-\alpha}\right) = \left(\mathrm{e}^{-\alpha}\right)^{n-1}\left(1 - \mathrm{e}^{-\alpha}\right) \end{aligned}$$

• Loi exponentielle et loi de Poisson. On suppose que le nombre N_t de clients se présentant à un guichet durant un intervalle de temps t suit la loi de Poisson $\mathcal{P}(\lambda t)$, $\lambda > 0$. Soit Y la v.a égale à l'intervalle de temps séparant l'arrivée de deux clients consécutifs. Y est à valeurs dans $\mathbf{R}^+$, et, pour tout $t > 0$, $(N_t = 0) = (Y > t)$, donc

$$\mathrm{P}(Y > t) = \mathrm{e}^{-\lambda t}\frac{(\lambda t)^0}{0!} = \mathrm{e}^{-\lambda t}\ ; \quad \mathrm{P}(Y \leqslant t) = 1 - \mathrm{e}^{-\lambda t}$$

Et comme $\mathrm{P}(Y \leqslant t) = 0$ si $t \leqslant 0$, on en déduit que $Y \hookrightarrow \mathcal{E}(\lambda)$.

2.3 Loi normale, ou loi de Laplace-Gauss

Elle modélise de très nombreuses séries statistiques (taille, poids des individus d'une population...). Voir aussi § 9.3.4, approximations.

Soit $m \in \mathbf{R}$ et $\sigma > 0$. On dit que X suit la **loi normale** de paramètres m, σ^2, et on note $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$, ssi X admet pour densité la fonction f telle que, pour tout $t \in \mathbf{R}$:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-m)^2}{2\sigma^2}}$$

Espérance, variance

$$\mathrm{E}(X) = m \;;\quad \mathrm{V}(X) = \sigma^2$$

On a montré au § 3.3.4 que l'intégrale de f sur $\mathbf{R}$ était convergente.

Loi normale centrée réduite

• Pour une v.a X quelconque admettant une espérance et une variance non nulle, la **v.a centrée réduite** associée X^* est définie par

$$X^* = \frac{X - \mathrm{E}(X)}{\sigma(X)}$$

On a $\mathrm{E}(X^*) = 0$; $\mathrm{V}(X^*) = 1$.

Proposition 1. $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$ si et seulement si la v.a centrée réduite X^* associée suit la loi normale $\mathcal{N}(0, 1)$, ou **loi normale centrée réduite** :

$$X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right) \Leftrightarrow X^* = \frac{X - m}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$$

• Soit $X^* \hookrightarrow \mathcal{N}(0, 1)$.

– X^* a pour densité la fonction φ définie par $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ et par conséquent (**intégrale de Gauss**) :

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t = \sqrt{2\pi}$$

– La fonction de répartition de X^* est notée $\lambda\Phi$. Pour tout $x \in \mathbf{R}$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t$$

– $\mathrm{E}(X^*) = 0$; $\mathrm{V}(X^*) = 1$.

Utilisation de la loi normale centrée réduite

• Grâce à la proposition 1, les calculs sur les v.a gaussiennes (i.e. suivant une loi normale) se ramènent à des calculs sur la loi normale centrée réduite. Plus précisément :

Proposition 2

• Soit $a, b, x \in \mathbf{R}$, avec $a < b$, et $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$. Alors

$$\mathrm{P}(X \leqslant x) = \Phi\left(\frac{x-m}{\sigma}\right)$$

$$\mathrm{P}(a \leqslant X \leqslant b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$$

• Pour tout $x \in \mathbf{R}$,

$$\Phi(-x) = 1 - \Phi(x).$$

• Les valeurs de Φ ; sont tabulées (uniquement pour les valeurs positives, ce qui est suffisant en utilisant le 2° de la proposition 2).

Pour toute v.a $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$:

$$\begin{aligned}\mathrm{P}\left(|X-m| \leqslant \sigma\right) &= \mathrm{P}(m-\sigma \leqslant X \leqslant m+\sigma) \\ &= \Phi\left(\frac{m+\sigma-m}{\sigma}\right) - \Phi\left(\frac{m-\sigma-m}{\sigma}\right) \\ &= \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 0{,}68\end{aligned}$$

• À partir de la loi normale centrée réduite, on peut retrouver la valeur de certaines intégrales.

$\displaystyle\int_0^{+\infty} t^2 e^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{\sqrt{2\pi}}{2}$. En effet, avec $X^* \hookrightarrow \mathcal{N}(0,1)$, on a $\mathrm{E}(X^*) = 0$ et $\mathrm{V}(X^*) = 1$, donc

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (t-0)^2 e^{-\frac{t^2}{2}}\,\mathrm{d}t = 1, \quad \text{donc} \quad \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}}\,\mathrm{d}t = \sqrt{2\pi}$$

D'où le résultat, car la fonction à intégrer est paire.

3. Convergences et approximations

3.1 Inégalité de Bienaymé - Tchebycheff

Théorème. Soit X une variable aléatoire discrète ou à densité, possédant une espérance $\mathrm{E}(X)$ et une variance $\mathrm{V}(X)$.
Alors :

$$\forall \varepsilon > 0, \quad \mathrm{P}\left(|X - \mathrm{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}$$

La probabilité qu'une v.a s' écarte de plus de ε de sa valeur moyenne est d'autant plus faible que sa variance est petite et que ε est grand.

De façon équivalente, on a :

$$\forall \varepsilon > 0, \quad \mathrm{P}\left(|X - \mathrm{E}(X)| < \varepsilon\right) \geqslant 1 - \frac{\mathrm{V}(X)}{\varepsilon^2}$$

3.2 Loi faible des grands nombres

Théorème. Soit (X_n) une suite de v.a indépendantes et de même loi, d'espérance m et de variance σ^2 positive.

Soit $Z_n = \dfrac{1}{n}(X_1 + X_2 + \cdots + X_n)$.

Alors, pour tout $\varepsilon > 0$:

$$\mathrm{P}\left(|Z_n - m| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

Il en résulte :

$$\lim_{n\to+\infty} \mathrm{P}\left(|Z_n - m| \geqslant \varepsilon\right) = 0 \ ; \qquad \lim_{n\to+\infty} \mathrm{P}\left(|Z_n - m| < \varepsilon\right) = 1$$

On dit que la suite de v.a (Z_n) **converge en probabilité** vers la variable aléatoire certaine m.

De façon générale, on dit que la suite de v.a (X_n) converge en probabilité vers la v.a X ssi :

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} \mathrm{P}\left(|X_n - X| \geqslant \varepsilon\right) = 0$$

Démonstration de la loi faible des grands nombres : on applique l'inégalité de Bienaymé-Tchebycheff à Z_n, dont l'espérance est égale à m, et la variance à $\frac{\sigma^2}{n}$.

3.3 Convergence en loi

Exemple 1

Théorème. Soit $\lambda > 0$ fixé, et soit, pour tout $n \in \mathbf{N}^*$, X_n une v.a de loi $\mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Alors

$$\forall k \in \mathbf{N}, \qquad \lim_{n \to +\infty} \mathrm{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

On dit que la suite de v.a (X_n) **converge en loi** vers une v.a de loi $\mathcal{P}(\lambda)$.
Ce théorème a été démontré § 8.4.2.

Exemple 2 : Théorème de la limite centrée

Théorème. Soit (X_n) une suite de v.a indépendantes et de même loi, d'espérance m et de variance σ^2 positive.
Soit S_n^* la variable centrée réduite associée à

$$S_n = X_1 + X_2 + ... + X_n$$

Alors $\left(S_n^*\right)$ converge en loi vers X^*de loi normale centrée réduite, c'est-à-dire que l'on a, pour tout a, b dans $\mathbf{R}$ tels que $a < b$:

$$\lim_{n \to +\infty} \mathrm{P}\left(a \leqslant S_n^* \leqslant b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} \, \mathrm{d}t$$

Théorème admis.

• $\mathrm{E}(S_n{}^*) = nm$; $\mathrm{V}(S_n{}^*) = n\sigma^2$; Le théorème de la limite centrée signifie donc, pour tout a, b dans $\mathbf{R}$ tels que $a < b$:

$$\lim_{n \to +\infty} \mathrm{P}\left(a \leqslant \frac{S_n - nm}{\sigma\sqrt{n}} \leqslant b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} \, \mathrm{d}t$$

• Soit $Z_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \frac{S_n}{n}$.
On a d'après l'égalité ci-dessus :

$$\lim_{n \to +\infty} \mathrm{P}\left(a \leqslant \frac{Z_n - m}{\frac{\sigma}{\sqrt{n}}} \leqslant b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} \, \mathrm{d}t$$

Et comme $\mathrm{E}(Z_n) = m$, $\mathrm{V}(Z_n) = \frac{\sigma^2}{n}$, on obtient :

$$\lim_{n \to +\infty} \mathrm{P}\left(a \leqslant Z_n^* \leqslant b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} \, \mathrm{d}t$$

Avec Z_n^* v.a centrée réduite associée à Z_n , la suite $\left(Z_n^*\right)$ converge elle aussi vers X^* de loi normale centrée réduite.

• De façon générale, on dit que la suite de v.a (X_n) converge en loi vers la v.a X ssi, avec F_n f.r de X_n et F f .r de X :

$$\lim_{n\to+\infty} F_n(x) = F(x)$$

en tout point x de $\mathbf{R}$ où F est continue, et on admet que cette définition générale se traduit comme indiqué dans les deux cas évoqués ici.

3.4 Approximations

• Si le taux de sondage $\frac{n}{N}$ est inférieur à $\frac{1}{10}$, la loi hypergéométrique $\mathcal{H}(N, n, p)$ peut être approchée par la loi binomiale $\mathcal{B}(n, p)$.

• Si n est grand et p petit, la loi binomiale $\mathcal{B}(n, p)$ peut être approchée par la loi de Poisson $\mathcal{P}(np)$.

• Si n est grand et p ni trop grand ni trop petit, la loi binomiale $\mathcal{B}(n, p)$ peut être approchée par la loi normale $\mathcal{N}(np, np(1-p))$.

• Si λ est grand, $\mathcal{P}(\lambda)$ peut être approchée par $\mathcal{N}(\lambda, \lambda)$.

On donne ici des critères qualitatifs (« n grand »...), il existe des critères plus précis.

Si n est grand et p petit, la loi binomiale $\mathcal{B}(n, p)$ peut être approchée par la loi de Poisson $\mathcal{P}(np)$: cela signifie que dans les conditions mentionnées, si $X \hookrightarrow \mathcal{B}(n, p)$ et $\widetilde{X} \hookrightarrow \mathcal{P}(np)$, pour tout $k \in [0, n]$, $\mathrm{P}\left(\widetilde{X} = k\right)$ constitue une bonne approximation de $\mathrm{P}(X = k)$. C'est une traduction de la convergence en loi de $X_n \hookrightarrow \mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$ vers $X \hookrightarrow \mathcal{P}(\lambda)$.

Dans le cas où on approche une v.a discrète par une v.a absolument continue, il faut procéder à la « correction de continuité ». Par exemple, avec λ grand et $X \hookrightarrow \mathcal{P}(\lambda)$:

$$\mathrm{P}(X = n) \approx \mathrm{P}\left(n - \frac{1}{2} \leqslant \widetilde{X} \leqslant n + \frac{1}{2}\right), \quad \text{avec} \quad \widetilde{X} \hookrightarrow \mathcal{N}(\lambda, \lambda)$$

Dans tous les cas, on approche une loi par une autre loi ayant même espérance, et même variance s'il y a besoin d'ajuster deux paramètres.

• Du théorème de la limite centrée, on infère le résultat suivant :

Soit (X_n) une suite de v.a indépendantes et de même loi, d'espérance m et de variance $\sigma^2 > 0$, et soit

$$S_n = X_1 + X_2 + \cdots + X_n \ ; \quad Z_n = \frac{S_n}{n}$$

Alors S_n suit approximativement la loi $\mathcal{N}\left(nm, n\sigma^2\right)$, et Z_n suit approximativement la loi $\mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$.

4. Estimation

Une *expérience aléatoire* donne lieu *à une observation numérique*. Cela définit une variable aléatoire X. Dans la pratique, la loi de X n'est pas complètement connue, elle dépend d'un paramètre que l'on cherche à estimer.

• On lance une pièce de monnaie. $X = 1$ si on obtient pile, 0 sinon.

X suit la loi de Bernoulli $\mathcal{B}(p)$. p (probabilité d'obtenir « pile ») est le paramètre à estimer.

• Des véhicules se présentent à un péage de faible trafic.

X, nombre de véhicules se présentant durant un intervalle de temps de 10 mn, suit la loi de Poisson $\mathcal{P}(\lambda)$. On cherche à estimer λ.

• Une photocopieuse tombe en panne. On peut chercher à modéliser sa durée de bon fonctionnement par une v.a X suivant la loi exponentielle $\mathcal{E}(\alpha)$. α est inconnu et on cherche à l'estimer.

4.1 Estimation ponctuelle

Définitions. Soit X la variable aléatoire non complètement connue, supposée avoir une espérance et une variance.

Soit θ le paramètre à estimer.

• Pour cela on réalise n fois l'expérience aléatoire.

On obtient un ***n*–échantillon** $(X_1, \cdots, X_n)$: $X_1, \cdots, X_n$ sont des variables aléatoires de même loi que X, et supposées indépendantes.

• Un **estimateur** de θ est une variable aléatoire T_n fonction uniquement de $X_1, X_2, \cdots, X_n$ et de n.

Problème : Apprécier quantitativement différents estimateurs, afin de choisir « le meilleur ».

• On appelle **biais** de T_n la quantité

$$b(T_n) = b = \mathrm{E}(T_n) - \theta.$$

T_n est dit **sans biais** ssi $\mathrm{E}(T_n) = \theta$.

• On appelle **risque quadratique** de T_n la quantité :

$$\mathrm{r}(T_n) = \mathrm{E}\left((T_n - \theta)^2\right)$$

Propriétés

La **moyenne empirique**

$$Z_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

est un estimateur sans biais de $\mathrm{E}(X)$.

En effet, d'après la linéarité de l'espérance :

$$\begin{aligned} E(Z_n) &= \frac{1}{n}(E(X_1) + E(X_2) + \cdots + E(X_n)) \\ &= \frac{1}{n} n\mathrm{E}(X) = \mathrm{E}(X) \end{aligned}$$

Avec $r(T_n)$ le risque quadratique et b le biais de T_n :

$$\mathrm{r}(T_n) = b^2 + \mathrm{V}(T_n)$$

Démonstration :

$$\begin{aligned} (T_n - \theta)^2 &= [T_n - \mathrm{E}(T_n) + \mathrm{E}(T_n) - \theta]^2 = [T_n - \mathrm{E}(T_n) + b]^2 \\ &= (T_n - \mathrm{E}(T_n))^2 + 2b(T_n - \mathrm{E}(T_n)) + b^2 \end{aligned}$$

Et on obtient le résultat en utilisant la linéarité de l'espérance. En effet, l'espérance de $(T_n - E(T_n))^2$ est par définition égale à la variance de T_n, et l'espérance de $T_n - \mathrm{E}(T_n)$ est égale à 0.

Remarques

- Entre deux estimateurs, on préférera souvent celui dont le risque quadratique est le plus petit.
- Pour un estimateur sans biais, le risque quadratique est égal à la variance.
- On parle aussi d'estimateurs *asymptotiquement sans biais* :

$$\lim_{n \to +\infty} \mathrm{b}(T_n) = 0$$

- Un estimateur est dit *convergent* ssi

$$\lim_{n \to +\infty} \mathrm{r}(T_n) = 0$$

Par exemple, la moyenne empirique est un estimateur convergent de l'espérance.

- En toute rigueur, l'espace probabilisable sous-jacent est muni, dans le cadre de l'estimation, d'une *famille de probabilités* P_θ dépendant du paramètre θ. Les notations $\mathrm{E}(T_n)\ \ldots$ sont donc des abus d'écriture, qui ne portent pas à conséquence.

4.2 Estimation par intervalle de confiance

Soit p une probabilité inconnue, ou le paramètre inconnu d'une loi de probabilité.

Soit α un nombre appartenant à $]0\,;1[$.

On dit que l'intervalle $[a\,;b]$ est un **intervalle de confiance** pour p **au niveau de confiance** $1 - \alpha$ (ou au **niveau de risque** α) ssi :

$$P(a \leqslant p \leqslant b) \geqslant 1 - \alpha$$

α est souvent donné sous forme de pourcentage : on parlera d'un intervalle de confiance au niveau de confiance 95 % par exemple .

On obtient un intervalle de confiance pour p en utilisant l'inégalité de Bienaymé-Tchebycheff, ou une approximation par une loi normale, *via* le théorème de la limite centrée.

On lance 1 000 fois une pièce de monnaie, on obtient 480 fois « pile ». La probabilité d'obtenir « pile » avec cette pièce est une inconnue p, et on cherche à l'estimer.

Pour i entier compris entre 1 et 1 000, on considère la variable aléatoire X_i définie par : $X_i = 1$ si on a obtenu « pile » au i-ème lancer, $X_i = 0$ sinon.

X_i suit la loi de Bernoulli de paramètre p. On a $E(X_i) = p$, $V(X_i) = p(1-p)$.

On considère

$$S_{1\,000} = \sum_{i=1}^{1\,000} X_i \;; \qquad Z_{1\,000} = \frac{S_{1\,000}}{1\,000} = \frac{1}{1\,000}\sum_{i=1}^{1\,000} X_i$$

- Appliquons l'inégalité de Bienaymé-Tchebycheff à la v.a $Z_{1\,000}$:

$$\forall \varepsilon > 0, P\left(|Z_{1\,000} - E(Z_{1\,000})| \geqslant \varepsilon\right) \leqslant \frac{V(Z_{1\,000})}{\varepsilon^2}$$

D'après la linéarité de l'espérance, $E(Z_{1\,000}) = p$, et, en utilisant la formule $V(aX+b) = a^2 V(X)$, puis l'indépendance des v.a X_i : $V(Z_{1\,000}) = \frac{p(1-p)}{1\,000}$. On obtient donc :

$$\forall \varepsilon > 0, P\left(|Z_{1\,000} - p| \geqslant \varepsilon\right) \leqslant \frac{p(1-p)}{1\,000\varepsilon^2}$$

Par conséquent :

$$\forall \varepsilon > 0, P\left(|Z_{1\,000} - p| < \varepsilon\right) \leqslant P\left(|Z_{1\,000} - p| \leqslant \varepsilon\right) \geqslant 1 - \frac{p(1-p)}{1\,000\varepsilon^2}$$

Pour tout $p \in [0\,;1]$, on a $0 \leqslant p(1-p) \leqslant \frac{1}{4}$ (étudier la fonction $f(t) = t(1-t)$ sur $[0\,;1]$). Donc

$$1 - \frac{p(1-p)}{1\,000\varepsilon^2} \geqslant 1 - \frac{1}{4\,000\varepsilon^2} \geqslant 0{,}90 \text{ pour tout } \varepsilon \geqslant \frac{1}{20}.$$

On a obtenu finalement

$$\mathrm{P}\left(|Z_{1\,000} - p| < \frac{1}{20}\right) \geqslant 0{,}90$$

Or la valeur observée de $Z_{1\,000}$ est $480/1\,000 = 0{,}48$, et

$$|0{,}48 - p| < \frac{1}{20} \Leftrightarrow -\frac{1}{20} < 0{,}48 - p < \frac{1}{20}$$
$$\Leftrightarrow 0{,}48 - \frac{1}{20} < p < 0{,}48 + \frac{1}{20}$$

Un intervalle de confiance pour p au niveau de confiance 0,90 est donc $]0{,}43\,;0{,}53[$.

• Maintenant, utilisons le théorème de la limite centrée, appliquée à $Z_n = \frac{1}{n}\sum_{i=0}^{n} X_i$: la variable centrée réduite associée à Z_n converge en loi vers une v.a T qui suit la loi normale centrée réduite.

1 000 est assez « grand », on peut donc considérer que

$$T = \frac{Z_{1\,000} - \mathrm{E}(Z_{1\,000})}{\sigma(Z_{1\,000})} = \frac{Z_{1\,000} - p}{\sqrt{\frac{p(1-p)}{1\,000}}}$$

suit la loi normale centrée réduite. Par lecture inverse de la table de la loi normale centrée réduite, on obtient

$$\mathrm{P}(-1{,}65 \leqslant T \leqslant 1{,}65) \geqslant 0{,}90$$

On a donc, en tenant compte de $p(1-p) \leqslant \frac{1}{4}$:

$$\mathrm{P}\left(-1{,}65 \times \frac{1}{2\sqrt{1\,000}} \leqslant 0{,}48 - p \leqslant 1{,}65 \times \frac{1}{2\sqrt{1\,000}}\right) \geqslant 0{,}90$$

On obtient finalement, par cette méthode

$$[0{,}453\ ;0{,}507]$$

comme intervalle de confiance pour p au niveau de confiance 0,90.



Partie 4

Informatique



Éléments d'algorithmique 10

1. Le langage PASCAL

1.1 Structure générale d'un programme

1°) En-tête du programme

Syntaxe	Commentaires
rogram nom ;	

2°) Déclaration des variables

var n:**integer** ;	n de type entier
x,moyenne:**real ;**	x et moyenne de type réel
liste:**array**[1..10]of integer ; mat:array[1..3,1..3]of real ;	liste et mat de type tableau ; liste[1], ... , liste[10] sont 10 variables de type entier ; mat[1,1], ... , mat[3,3] sont 9 variables de type réel
test: **boolean** ;	test de type booléen ; sa valeur est **true** (vrai) ou **false** (faux)

3°) Déclaration (éventuelle) des procédures et fonctions : voir § 10.1.8.

4°) Programme principal

BEGIN < instructions > **END.**	une instruction se termine par un point-virgule, le programme principal par un point final

Cette structure doit être scrupuleusement respectée ; dans l'ordre :

- **1°), 2°), 4°)** pour un programme simple ; 1°), 2°) est la **partie déclarative** du programme, 4°) en est le **corps**.
- **1°), 2°), 3°), 4°)** pour un programme structuré.

1.2 Variables

Une **variable** au sens informatique est un emplacement dans la mémoire de l'ordinateur. Cet emplacement est créé quand la variable est déclarée.

> Si une variable `n` de type `integer` est déclarée, un emplacement de 2 octets est réservé. Cet emplacement est destiné à recevoir une certaine valeur entière appartenant à $[\![-32768, 32767]\!]$, appelée le **contenu** (ou valeur) de la variable `n`. Ce contenu peut changer au cours de l'exécution du programme.

Un **algorithme** informatique est une suite d'instructions destinée à produire un certain contenu pour une ou plusieurs **variables de sortie**, à partir du contenu d'une ou plusieurs **variables d'entrée**.

> **Initialiser** une variable, c'est lui donner une première valeur.
> Une variable doit toujours être initialisée.

On initialise une variable d'entrée par **lecture** :

`write``('n =') ;`	instruction facultative, qui affiche à l'écran les caractères entre comas (') ; l'utilisateur entre alors la valeur souhaitée au clavier
`readln``(n) ;`	la variable `n` reçoit cette valeur comme contenu

On initialise une variable de sortie par **affectation** :

`x := 0 ;`	
`moyenne := (a+b) 2 ;`	`a` et `b` ont été initialisés
`test := false ;`	
`for i := 1 to 10 do`	`liste[1]` reçoit la valeur 10,
`liste[i]:= 11-i ;`	...
	`liste[10]` reçoit la valeur 1

Quand les variables de sortie ont le contenu souhaité, il faut afficher à l'écran ce contenu :

`write(a,b) ;`	affichage des **contenus** de `a`, `b`. Ne pas confondre avec : **`write('a,b') ;`** , qui affiche exactement « a,b »
`writeln(a,b) ;`	idem, puis passage à la ligne

1.3 Opérateurs et fonctions

Opérateurs et fonctions arithmétiques

opérateurs	entrée(s)	sortie	commentaire
+ - *	réel/ entier	réel/entier	***** : multiplication
	réel ou entier	réel	division
ln ex sqrt	réel ou entier	réel	sqrt : racine carrée
trunc	réel	entier	partie entière
abs	réel/entier	réel/entier	valeur absolue

Mentionnons aussi **div** et **mod** , quotient et reste dans la division euclidienne. Entrées et sorties sont de type entier.

> Le contenu de 23 div 5 est **4**, le contenu de 23 mod 5 est **3**, car $23 = 5 \times \mathbf{4} + \mathbf{3}$.

Opérateurs logiques

Entrée(s) et sortie sont de type booléen (reçoivent la valeur true ou false).

	not
T	F
F	T

	q	**or** q	**and** q
T	T	T	T
T	F	T	F
F	T	T	F
F	F	F	F

Opérateurs relationnels

Les variables d'entrée sont de type réel ou entier. La variable de sortie est de type booléen :

= ; **<>** pour $\neq$; **<** ; **>** ; **<=** pour $\leqslant$; **>=** pour $\geqslant$

> La variable (a = b) a le contenu true si le contenu de a est égal au contenu de b, et le contenu false si le contenu de a n'est pas égal au contenu de b.

Les opérateurs relationnels sont utilisés principalement dans les instructions conditionnelles :

if < relation > then < instructions > [else < instructions >]

while < relation > do < instructions >

re eat < instructions > until < relation >

1.4 Générateurs de nombres aléatoires

• On écrit l'instruction **randomize** juste après le BEGIN du programme principal.

• Le contenu de **random** est un nombre réel (pseudo-) aléatoire appartenant à [0 ; 1], suivant la loi uniforme sur cet intervalle.

random est de type réel. Voir § 9.2.1.

> Affichage d'un nombre aléatoire appartenant à [0, 1] :
> ```
>  rogram alea ;BEGIN randomize ;write(random) ;END.
> ```

• Si n est un entier $\geqslant 0$, le contenu de random(n) est un nombre entier (pseudo-) aléatoire appartenant à l'ensemble $\{0, \ldots, n-1\}$, en suivant la loi uniforme sur cet ensemble (voir § 7.4.4).

n peut être éventuellement une variable de type entier, initialisée.

random(n) est de type entier.

> Simulation du lancer d'un dé cubique équilibré :
> ```
>  rogram de ;
> var lancer:integer ;
> BEGIN randomize ;
> lancer := random(6)+1 ;writeln(lancer) ;END.
> ```

1.5 Instruction conditionnelle *IF THEN [ELSE]*

if <relation> **then** begin < instructions A > end	Si relation a la valeur true, les instructions A sont exécutées ; sinon les instructions B sont exécutées. Avec une seule instruction A, le begin else est inutile, idem pour les instructions B.
else begin < instructions B > end ;	else ne doit pas être précédé immédiatement d'un point virgule. Le else et les instructions B sont optionnelles.

1.6 Boucles définies

FOR i := n1 **TO** n2 **DO** begin < instructions A > end ;	`n1` et `n2` sont deux nombres entiers, ou deux variables initialisées de type entier. Les instructions A sont exécutées avec `i` de contenu `n1`, puis avec `i` de contenu `n1+1`,. . . , puis une dernière fois avec `i` de contenu `n2`.

Si `n1>n2`, les instructions A ne sont pas exécutées.

Variante :

FOR i := n1 **DOWNTO** n2 **DO** begin < instructions A > end ;	Les instructions A sont exécutées avec `i` de contenu `n1`, puis avec `i` de contenu `n1 - 1`,. . . , puis une dernière fois avec `i` de contenu `n2`.

Si `n1<n2`, les instructions A ne sont pas exécutées.

Chacun des passages dans la boucle, c'est-à-dire chacune des exécutions du groupe d'instructions A, est une **itération**.

Comme dans le cas de l'instruction conditionnelle, si il n'y a qu'une seule instruction A, le `begin end ;` est inutile.

1.7 Boucles conditionnelles

REPEAT < instructions > **UNTIL** < relation >	Les instructions délimitées par le `re eat until` sont exécutées jusqu'à ce que la relation devienne vraie. Elles sont exécutées au moins une fois.

La relation doit devenir vraie, sinon on entre dans une boucle infinie.

WHILE < relation > **DO** begin < instructions > end ;	Les relations délimitées par le `begin end ;` sont exécutées tant que la relation est vraie. Elles peuvent ne pas être du tout exécutées.

La relation doit devenir fausse, sinon on entre dans une boucle infinie.

Si il y a une seule instruction à exécuter, le `begin end ;` est inutile.

1.8 Procédures et fonctions

Un exemple de procédure dans un programme structuré

Le programme suivant classe deux nombres réels donnés par l'utilisateur en ordre croissant. Les commentaires sont mis entre accolades, une possibilité qu'offre le langage PASCAL.

```
rogram tri ;
{en-tête du programme}
var a,b :real ;
{déclaration des variables du programme, dites globales}
{--------------------------------------------------}
rocedure echanger(VAR x,y:real) ;
{en-tête de la procédure, avec x et y comme paramètres}
var aux:real ;
{déclaration des variables internes à la procédure, dites locales}
begin
  aux:= x ; x:= y ; y:= aux ; writeln(x,y) ;
{avant d'effectuer x :=y, on stocke le contenu de x dans une variable
auxiliaire, sinon sont contenu est perdu}
end ;{le begin end encadre le corps de la procédure}
{--------------------------------------------------}
BEGIN
{début du programme principal}
write('a=') ; readln(a) ;
write('b=') ;readln(b) ;
{initialisation des variables d'entrée}
if b>a then echanger(a,b) ;
{si b > a, il y a appel de la procédure echanger, qui est alors exécutée
avec les paramètres x,y remplacés par les variables a,b}
writeln(a,b) ;
{affichage du résultat}
END. {fin du programme principal}
```

On voit sur cet exemple qu'on définit une procédure de la même manière qu'on définit un programme simple :

- partie déclarative :
  – en-tête de la procédure, avec ses éventuels paramètres ;
  – déclaration des éventuelles variables locales ;
- corps de la procédure.

Passage de paramètres en valeur et en variable

• Reprenons l'en-tête de la procédure `echanger` dans le programme `tri` précédent :

```
rocedure echanger(VAR x,y : real) ;
```

Si l'utilisateur entre les nombres : 2 ; 19 ; il obtiendra à l'écran

19 2 (écriture dans le corps de la procédure)

19 2 (écriture dans le programme principal)

Le passage des paramètres se fait **en variable** : mot-clé `VAR` devant les paramètres. La valeur des variables `a`, `b` est affectée par l'exécution de la procédure.

• Si on remplace l'en-tête de la procédure `echanger` dans le même programme par :

```
rocedure echanger(x,y : real) ;
```

l'utilisateur qui entre les nombres 2 ; 19 obtiendra maintenant :

19 2 (écriture dans le corps de la procédure)

2 19 (écriture dans le programme principal)

Le passage des paramètres se fait **en valeur** : le mot clé `VAR` n'est plus présent devant les paramètres. La valeur des variables `a`, `b` n'est pas affectée par l'exécution de la procédure.

• Règle : on passe en valeur les paramètres correspondant à des variables d'entrée. On passe en variable les paramètres correspondant à des variables de sortie.

Un exemple de fonction

On définit la fonction `uisiter`, destinée à calculer x^n pour $x \in \mathbf{R}$ et $n \in \mathbf{N}^*$:

```
function  uisiter(x:real ;n:integer):real ;
{en-tête de la fonction}
var x n:real ;i:integer ; {variables locales}
begin
  x n :=1 ;
  for i :=1 to n do x n := x n*x ;
{à chaque passage dans la boucle, le contenu de x n est multiplié par x,
le résultat devient le nouveau contenu de x n}
   uisiter := x n ;
{il y a eu n passages dans la boucle, le contenu de x n est celui souhaité,
on l'affecte à  uisiter } ;
end ;{fin de la définition de la fonction}
```

Récursivité

Dans le corps de la fonction uisiter précédente, l'introduction de la variable locale xpn est artificiel, mais obligatoire pour une fonction définie de façon itérative.

Une autre possibilité est de définir une fonction uisrec de façon récursive, très proche de la définition mathématique du nombre x^n :

$$x^0 = 1$$

$$\text{si } n \geqslant 1, \ x^n = x^{n-1} \times x$$

valable pour tout $n \in \mathbf{N}$ et tout $x \in \mathbf{R}$, avec la convention $0^0 = 1$.

On écrit :

```
function  uisrec(x:real ;n:integer):real ;
begin
  if n=0 then  uisrec := 1
  else  uisrec :=  uisrec(n-1)*x ;
end ;
```

Quelques remarques

• On définit une fonction de la même de la même manière qu'on définit une procédure ou un programme simple :

– partie déclarative :

∗ en-tête de la fonction, avec ses éventuels paramètres ;

∗ déclaration des éventuelles variables locales ;

– corps de la fonction.

• Réciproquement, un programme simple peut toujours être réécrit pour obtenir une procédure ou une fonction suivant le cas (une fonction si la sortie est une variable de type entier, réel, tableau ou booléen ; une procédure si la sortie est plus complexe).

• Pour une fonction, on fait toujours passer les paramètres en valeur (absence du mot-clé `VAR` devant les paramètres) ;

• L'instruction `writeln ;` est une commodité d'affichage : elle produit à l'écran une ligne vide.

• Le langage PASCAL n'est pas sensible à la casse : il ne distingue pas majuscule et minuscule. On peut utiliser cette propriété pour mettre en valeur certaines instructions. Ici on a choisi de mettre en valeur le `BEGIN END` du programme principal et le mot-clé `VAR` devant les paramètres de sortie d'une procédure.

2. Exemples d'algorithmes

2.1 Algorithmes de base

Programmation d'une suite $u_{n+1} = f(u_n)$

• Calcul et affichage des 100 premiers termes de la suite (u_n) définie par

$$\begin{cases} u_0 = 1 \\ \forall n \in \mathbf{N}, u_{n+1} = f(u_n) \end{cases} \quad \text{avec } f(x) = e^{-x^2}$$

L'instruction cruciale est `u:=f(u) ;`

```
 rogram suitrec ;
{en-tête du programme}
var u:real ;n:integer ;
{déclaration des variables globales}
function f(x:real):real ;
begin
  f:= ex (-x*x) ;
end ;
{déclaration de la fonction f}
BEGIN
{début du programme principal}
n:=0 ;u:=1 ;
{initialisation des variables}
writeln(n,u) ;
{affichage du terme de rang 0 de la suite, avec son rang}
for n :=1 to 99 do
  begin
    u :=f(u) ;
    {le contenu u_n de u est remplacé par f(u_n) = u_{n+1}}
    writeln(n,u) ;
    {affichage du terme de rang n de la suite, avec son rang}
  end ;
END.
{fin du programme principal}
```

Remarque. Dans ce programme, `u` est une variable de sortie. La même variable peut être d'abord variable d'entrée, si la valeur de u_0 est donnée par l'utilisateur : `write('u_0=') ;readln(u) ;` .

Programmation d'une somme

L'instruction cruciale prend la forme

```
somme:=somme+nombre ;
```

• Calcul et affichage des des nombres

n ; u_n ; $S_n = \sum_{k=1}^{n} u_k$; avec $n \in \{1, \ldots, 100\}$, $u_n = f(n), f(x) = \frac{1}{x}$

```
 rogram somme ;
var n:integer ; S,u:real ;
function f(x:real):real ;
begin f:=1 x ;end ;
BEGIN
S:=0 ;
for n:=1 to 100 do
  begin
    u:=f(n) ;
    S:=S+u ;
    writeln(n,u,S) ;
  end ;
END.
```

On pourra adapter ce programme à d'autres fonctions f, par exemple :

$$f(x) = \frac{1}{x^2} \quad ; \quad f(x) = \frac{1}{x(x-1)} \quad ; \quad f(x) = \frac{1}{x^2-1} \quad ; \quad f(x) = \frac{1}{x^3}$$

• La série de terme général $\frac{1}{n}$, $n \in \mathbf{N}^*$ est divergente (voir § 2.4.3) : approche numérique. A étant un réel positif donné par l'utilisateur, on calcule la somme partielle $S_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ jusqu'à ce que sa valeur dépasse A, on affiche alors la valeur de S_n et le rang n.

```
 rogram harmo ;
var A,S:real ;n:integer ;
BEGIN
readln(A) ;{donner une valeur raisonnable pour A, 8 par exemple}
S:=0 ;n:=0 ;
re eat
  n:=n+1 ; S:=S+1 n ;
until S>A ;
writeln(S,n) ;
END.
```

2.2 Exemples en analyse

Calcul de $S_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$, première méthode

Les calculs de x^k et $k!$ sont traités par les fonctions uisrec et facrec, puis ces fonctions sont appelées dans la fonction S1.

Pour abréger, on a écrit des algorithmes récursifs, mais des versions itératives sont possibles, voir § 10.1.8 et le § 4.1 de l'Introduction.

```
function  uisrec(x:real ;n:integer):real ;
begin
  if n=0 then  uisrec := 1
  else  uisrec :=  uisrec(n-1)*x ;
end ;
function facrec(n :real):real ;
begin
  if n=0 then facrec :=1 else
  facrec := n*facrec(n-1) ;
end ;
function S1(x:real ;n:integer):real ;
begin if n=0 then S1:=1
    { avec n = 0, le contenu de S1 doit être x^0/0! = 1 }
else S1:=S1(x,n-1)+  uisrec(x,n) (facrec(x,n) ;
end ;
```

On constate par exemple que l'instruction writeln(S1(1,10)) ; dans un programme principal produit l'affichage d'une valeur approchée de e à moins de 10^{-6} près, ce qui est conforme à la théorie, voir § 2.4.2.

Calcul de $S_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$, deuxième méthode

– L'algorithme précédent est maladroit, car chaque appel des fonctions uisrec et facrec recommence tous les calculs au début, sans utiliser les relations de récurrence

$$x^{n+1} = x^n \times x, (n+1)! = n! \times n.$$

– On écrit une fonction S2 itérative. À chaque itération, on effectue Somme:=Somme+ , variable qui contient les valeurs successives

$$1 = \frac{x^0}{0!} \quad ; \quad \frac{x^1}{1!} \quad ; \quad \frac{x^2}{2!} \quad ; \quad \ldots$$

```
function S2(x:real ;n:integer):real ;
var Somme, :real ; k:integer ;
begin
 :=1 ; Somme:=1 ; {valeurs de x^0/0! et ∑_{k=0}^{0} x^k/k!}
for k:= 1 to n do
  begin
     := *x k ;
    {le contenu de  passe de x^(k−1)/(k−1)! à x^k/k!}
    Somme:=Somme+ ;
    {le contenu de Somme passe de S_{k−1}(x) à S_k(x)}
  end ;
S2:=Somme ;
end ;
```

Comparaison des deux méthodes

- Chaque appel récursif de la fonction S1

```
S1:=S1(x,n-1)+ uisrec(x,n) (facrec(x,n) ;
```

produit une addtion et une division et il y a `n` tels appels. Les appels des fonctions `uisrec(x,n)` et `facrec(x,n)` produisent eux-même, pour toute valeur de `n`, chacun `n-1` multiplications. On compte donc en tout :

`n` additions ; `n` divisions ;

`2(0 + 1 + 2 + ... + n-1) = n(n-1)` multiplications.

- Pour S2, chaque passage dans la boucle

```
for k:= 1 to n do
```

produit 1 multiplication, 1 division, 1 addition. On compte en tout

`n` additions ; `n` divisions ; `n` multiplications.

Soit un gain appréciable de temps ou de place (penser à `n` = 100).

2.3 Exemples en probabilités

Simulation d'une variable aléatoire X suivant la loi géométrique de paramètre $p \in]0, 1[$

- On écrit la procédure geom, d'en-tête

```
geom( :real ;VAR X:integer) ;
```

est le paramètre d'entrée (la probabilité du succès), `X` est le paramètre de sortie (le rang du premier succès).

```
 rocedure geom( :real ;VAR X:integer) ;
var resultat:real ;
begin
X:=0 ;
re eat
  resultat:= random ;
    {resultat est un réel aléatoire appartenant à [0 ;1]}
  if resultat <=  then write('S')
  else write('E') ;
    {la probabilité de resultat <=   est p, probabilité du succès}
  X:=X+1 ;
    {numéro du tirage en cours}
until resultat <=  ;
    {on continue jusqu'au succès}
Writeln(X) ;
end ;
```

• On peut utiliser la procédure ainsi définie dans un programme principal, par exemple sous la forme :

```
BEGIN
randomize ;
geom(0.1,Y) ;
writeln(Y) ;
END.
```

La variable `Y` est une variable (déclarée) du programme principal.

L'exécution du programme produira par exemple l'affichage

EEEEEES7 (écriture dans le corps de la procédure)

7 (écriture dans le programme principal)

Estimation de l'espérance d'une v.a

• On sait que si $(X_1, \ldots, X_n)$ est un n-échantillon de la v.a X, la moyenne empirique $Z_n = \frac{1}{n}(X_1 + \cdots + X_n)$ est un estimateur sans biais de l'espérance $\mathrm{E}(X)$, si celle-ci existe. Voir § 9.4.1.

• Soit X une v.a suivant la loi géométrique de paramètre p (pseudo-) inconnu. On utilise la procédure

```
geom( :real ;VAR X:integer) ;
```

définie précédemment pour donner une estimation de l'espérance

$\mathrm{E}(X) = \frac{1}{p}$ de cette variable aléatoire. On écrit uniquement le programme principal :

```
BEGIN
randomize ;
 :=random ; {le paramètre « inconnu » de X}
S:=0 ; {initialisation de la somme X1 + ··· + Xn }
for k:=1 to 1000 do {1000 taille de l'échantillon}
begin
  geom( ,X) ;
  S:=S+X ;
end ;
writeln(S 1000) ; {valeur donnée par l'estimateur Z1000}
writeln(1  ) ; {on compare à la valeur de l'espérance}
END.
```

Simulation d'une variable aléatoire suivant une loi hypergéométrique

• On considère une urne de taille N, contenant au départ b boules blanches, et de laquelle on tire une à une et sans remise nb boules. Le nombre de boules blanches obtenues suit la loi $\mathcal{H}\left(N, nb, b/N\right)$.

• On écrit une procédure récursive. L'idée générale de la récursivité est de se ramener de proche en proche à une problème simple. Le problème simple est ici le cas $nb = 1$.

• Les paramètres d'entrée sont `N, nb, b`. Le paramètre de sortie est `X`.

```
 rocedure hy erg(N,nb,b:integer ;VAR X:integer) ;
begin
  if nb=1 then if random<=b N then X:=X+1 ;
  if nb>1 then if random<=b N
    then
      begin
        X:=X+1 ;
        hy erg(N-1,nb-1,b-1,X) ;
      end
    else hy erg(N-1,nb-1,b,X) ;
end ;
```

• On initialisera la variable de sortie à 0 avant l'appel de la procédure hy erg dans un programme principal.

2.4 Exemples de gestion de listes

Recherche du maximum d'une liste de nombres entiers différents, et de son rang dans la liste

Les variables d'entrée sont n (taille de la liste) et t (la liste des n nombres, variable de type tableau) ; les variables de sortie sont

max et rangdumax.

```
 rogram maximum ;
var t :array[1..100] of integer ;
i,n,max,rangdumax:integer ;
BEGIN
write('n=') ;readln(n) ;
write('écrire',n,'nombres entiers différents') ;
for i :=1 to n do readln(t[i]) ;
max:=t[1] ;rangdumax:=1 ;
for i:=2 to n do
  begin
    if t[i]>max then
      begin
        max:=t[i] ;rangdumax:=i ;
      end ;
  end ;
writeln(rangdumax,max) ;
END.
```

Recherche dichotomique du rang c d'un élément donné X dans un liste donnée t de nombres entiers ordonnés par ordre croissant.

La récursivité est bien adaptée. Le problème simple auquel on se ramène est ici le cas t[c]=X.

```
function rang(a,b,X :integer) :integer ;
var c:integer ;
```

```
begin
c:=trunc((a+b) 2) ; {trunc partie entière}
if t[c]=X then rang:=c ;
if t[c]<X then rang:=rang(c,b,X) ;
if t[c]>x then rang:=rang(a,c,X) ;
end ;
```

On fera appel de cette fonction dans un programme principal de la manière suivante : `writeln(rang(1,n,X)) ;`
avec `t` et `n` déclarés et initialisés comme dans le programme précédent, et `X` déclaré de type entier et initialisé par lecture.





Index

S

T

V

LES RÉSUMÉS DU COURS



Gabriel Baudrand

MATHÉMATIQUES :

résumés du cours ECE 1re et 2e années

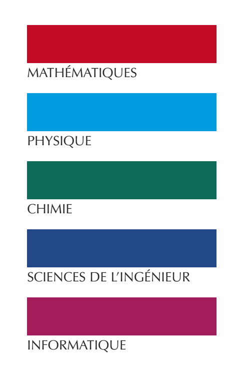
Préparez efficacement les devoirs, les interrogations orales et les concours !

Avec les résumés du cours de mathématiques, **révisez en un clin d'oeil les notions incontournables de première et de deuxième années**.

- **Tous les théorèmes, définitions et formules** du programme de mathématiques sur les deux années.
- Des exemples détaillés sur les applications clés
- Des **conseils**, des rappels de **méthode**, les **erreurs à éviter**.

GABRIEL BAUDRAND

est professeur agrégé de mathématiques en classes préparatoires au lycée Madeleine Michelis (Amiens).





ISBN 978-2-10-053972-7

www.dunod.com